The Dissertation Committee for Lakshay Narula
certifies that this is the approved version of the following dissertation:

Towards Secure & Robust PNT for Automated Systems

Committee:

Todd E. Humphreys, Supervisor

Haris Vikalo

Mohit Tiwari

Nuria Gonzalez-Prelcic

Brandon A. Jones

Towards Secure & Robust PNT for Automated Systems

by

Lakshay Narula

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2020

To the pursuit of knowledge and engineering.

Acknowledgments

I would like to express my sincere thanks and gratitude to all the people who have helped me bring this dissertation to bear.

I am thankful to my advisor, Prof. Todd Humphreys, for his consistent guidance and direction. In addition to my academic enrichment, I have always looked up to Todd for his technical writing, personnel management, generosity, and self-discipline. He always made sure that I received the recognition for our good work, while also having my back for whenever I made mistakes. Overall, I am grateful for his care and attention.

Very many thanks to my Ph.D. committee members Prof. Haris Vikalo, Prof. Mohit Tiwari, Prof. Nuria Gonzalez-Prelcic, and Prof. Brandon Jones for reviewing my Ph.D. research and serving on my dissertation committee.

The work in this dissertation received financial support from the Oak Ridge National Lab, the U.S. Department of Transportation, Samsung, Honda, and Toyota. I would like to acknowledge these agencies for their generous funding.

I would like to express my appreciation to my colleagues at the UT Radionavigation Lab for always holding ourselves to the highest standards. Technical discussions with Matthew Murrian, Michael Wooten, Peter Iannucci, Tucker Haydon, Daniel LaChapelle, and Nick Montalbano have been

very insightful and rewarding. Many parts of this dissertation started off in collaboration with Matthew, Michael, and Peter.

I would not be in a position to write this dissertation if not for the love and support of my family and friends. I thank my parents for the many unsaid sacrifices you have made over the years to support my aspirations. I also thank my sister and my brother for their support and counsel. Thanks to my friend Kanchan for accompanying me through the doctoral journey from the first day to the last, and to my friends Tucker, Michael, Caleb, Rebal, and Romir for all the good times.

Hook 'em Horns.

Towards Secure & Robust PNT for Automated Systems

Lakshay Narula, Ph.D.
The University of Texas at Austin, 2020

Supervisor: Todd E. Humphreys

This dissertation makes four contributions in support of secure and robust position, navigation, and timing (PNT) for automated systems. The first two relate to PNT security while the latter two address robust positioning for automated ground vehicles.

The first contribution is a fundamental theory for provably-secure clock synchronization between two agents in a distributed automated system. All one-way synchronization protocols, such as those based on the Global Positioning System (GPS) and other Global Navigation Satellite Systems (GNSS), are shown to be vulnerable to man-in-the-middle delay attacks. This contribution is the first to identify the necessary and sufficient conditions for provably secure clock synchronization.

The second contribution, also related to PNT security, is a three-year study of the world-wide GPS interference landscape based on data from a dual-frequency GNSS receiver operating continuously on the International Space

Station (ISS). This work is the first publicly-reported space-based survey of GNSS interference, and unveils previously-unreported GNSS interference activity.

The third contribution is a novel ground vehicle positioning technique that is robust to GNSS signal blockage, poor lighting conditions, and adverse weather events such as heavy rain and dense fog. The technique relies on sensors that are commonly available on automated vehicles and are insensitive to lighting and inclement weather: automotive radar, low-cost inertial measurement units (IMUs), and GNSS. Remarkably, it is shown that, given a prior radar map, the proposed technique operating on data from off-the-shelf all-weather automotive sensors can maintain sub-50-cm horizontal position accuracy during 60 min of GNSS-denied driving in downtown Austin, TX.

This dissertation's final contribution is an analysis and demonstration of the feasibility of crowd-sourced digital mapping for automated vehicles. Localization techniques, such as the one described in the previous contribution, rely on such digital maps for accuracy and robustness. A key enabler for large-scale up-to-date maps is enlisting the help of the very consumer vehicles that need the map to build and update it. A method for fusing multi-session vision data into a unified digital map is developed. The asymptotic limit of such a map's globally-referenced position accuracy is explored for the case in which the mapping agents rely on low-cost GNSS receivers performing standard code-phase-based navigation. Experimental validation along a semi-urban route shows that low-cost consumer vehicles incrementally tighten the accuracy of

the jointly-optimized digital map over time enough to support sub-lane-level positioning in a global frame of reference.

Table of Contents

List of Tables

List of Figures

Part I

Introduction

Chapter 1

Introduction

As critical infrastructure and safety-of-life systems become ever more automated, the security and robustness of such systems must be thoroughly examined. It is paramount that these systems have acceptable behavior under naturally-occurring challenging inputs (e.g., outliers and faults) as well as under attacks (e.g., outlaws and frauds). This dissertation focuses on the position, navigation, and timing (PNT) component of such automated systems, making fundamental contributions to PNT security and robustness.

Since the turn of the 21st century, the global positioning system (GPS) and other such global navigation satellite systems (GNSS) have become the *de facto* choice for positioning and timing solutions in critical infrastructure world-wide, so much so that GPS is often called an "invisible utility." Transportation, agriculture, energy, finance, and telecommunication all rely on GNSS-derived PNT for everyday operation [16]. Yet the vulnerability of civil GNSS receivers to counterfeiting-type attacks, commonly referred to as *spoofing*, is well-documented [46, 66, 71, 123]. Similarly, GNSS jamming amounts to a denial-of-service-type attack. While a jamming attack does not lead to misleading information, it is much simpler to execute, is far more common in-

the-wild [60, 106, 159]. GNSS impairment in the face of spoofing or jamming is often a single point-of-failure for large automated systems.

As regulatory bodies catch up with this growing threat, there have been calls to replace or augment GNSS [31, 114]. But how can one ensure that alternatives or augmentations to GNSS are not themselves vulnerable to cyber-physical attacks? Formalizing even the concept of cyber-physical system security is notoriously hard, to say nothing of provably establishing that a given system is secure. Ironically, it is often easier to prove the opposite: that a system is fundamentally vulnerable to a given attack.

Aside from counterfeit attacks, PNT solutions for critical infrastructure and safety-of-life systems must also deal with challenging natural phenomenon. Automated ground vehicles (AGVs), for example, are safety-of-life systems that rely on robust positioning and navigation for even the most basic tasks. The positioning engine, typically realized with a combination of GNSS and multiple automotive sensors, must provide reliable sub-lane-level accuracy with high availability in a variety of challenging environments. GNSS, for example, suffers from limited availability and accuracy in deep urban canyons due to signal blockage and multipath. Dead reckoning sensors like inertial measurement units (IMU) and wheel encoders can quickly drift out of acceptable accuracy limits unless corrected by other sensors [58]. Cameras are less useful in poor lighting, and both cameras and lidar perform poorly under adverse weather conditions such as heavy rain, dense fog, or a snowy whiteout. The real world is fraught with PNT edge cases that must be handled before

automated safety-of-life systems can be deployed en masse.

The goal of designing a robust and secure PNT service is a lofty one. This dissertation is certainly not the first to address this goal, and will not be the last. In the long struggle against outliers and outlaws, this dissertation seeks victories in a few key battles.

1.1 Towards Secure PNT

1.1.1 Provably Secure Clock Synchronization

This dissertation's first topic is secure clock synchronization. Attacks on clock synchronization (or time transfer) services can have serious implications for critical infrastructure such as smart power grids, telecommunication networks, financial markets, etc. [16]. Smart power grids demand global synchronization of their phasor measurement units (PMUs) with an accuracy of 26.5 µs [57, 92], which is much smaller than the period of the grids' alternating current oscillations. This synchronization is currently accomplished, whether directly or indirectly, using GNSS. A timing attack against key nodes in a power distribution network could disrupt the grid [139, 140]. A large-scale power outage due to a cyber attack has long been feared in the United States [61].

Like power grids, telecommunication networks employ GNSS for reliable time synchronization. 5G New Radio requires absolute timing synchronization accurate to 1.5 µs for efficient time division duplex operation [52]. With increased urban densification of cellular base stations, the telecommuni-

cation industry has expressed concerns that GNSS outages will degrade cellular performance. Meanwhile, purposeful attacks on clock synchronization could be far more catastrophic than intermittent outages.

Finally, competition between global financial exchanges for high-frequency traders, who are particularly concerned about measuring trading latency, have pushed exchanges toward millisecond-accurate timing or better [134]. Indeed, traders now even consider relativistic effects of their trades [7, 167]. Manipulation of exchange and market participant timing via attacks on clock synchronization could lead to confusion in the markets or illicit financial gains [146].

Despite these high-profile concerns, the security of clock synchronization methods against attacks has received little scrutiny in the literature. Unlike most cyber-physical attacks, it turns out that the problem of secure clock synchronization can be formally analyzed with a set of mathematical definitions, assumptions, and proofs. This dissertation is the first to make this observation and to establish a fundamental theory of provably secure clock synchronization, presented in Chap. 2. This work, published in [108], makes the following contributions:

- An argument that synchronization based on one-way communication between the master and slave clocks, e.g., with GNSS, cannot be provably secured.
- A proof that a proposed set of necessary and sufficient security conditions holds for a generic two-way synchronization system model.

- An analysis of specific timing protocols such as the IEEE 1588 Precision Time Protocol (PTP), leading to the conclusion that none of the major synchronization protocols currently in use is provably secure.

1.1.2 GNSS Interference Monitoring from Low-Earth Orbit

The threat of GNSS spoofing is no longer just an academic concern. Almost all civil GNSS receivers on the market today can be spoofed with cheap off-the-shelf radio equipment and open-source software. GNSS jamming, for its part, is also a security threat, as it can be used to selectively deny GNSS service. While a jamming attack does not lead to misleading information, it is much simpler to execute, is far more common in-the-wild [60,106,159], and can impact the security of other systems that rely on GNSS for positioning and/or timing, e.g., the TESLA broadcast authentication protocol [117]. Terrestrial GNSS interference activity "in the wild" has grown more widespread and sophisticated over recent years. Conspicuous GNSS jamming or spoofing has occurred, or is ongoing, at urban and coastal sites around the globe [3,27,30,137].

The second topic addressed in this dissertation is that of GNSS interference monitoring from low-earth orbit (LEO) satellites. The most obvious advantages of space-based interference observation from LEO are world-wide coverage and frequent flyovers at low altitude. No prior public literature explores the use of a space-borne GNSS receiver for monitoring terrestrial GNSS interference. Chap. 3 presents a three-year study of GNSS interference monitoring with a software-defined GNSS receiver aboard the International Space

Station (ISS). This work constitutes one part of the work published in [106]. The author of this dissertation made the following contributions to [106]:

- An analysis of the expected performance for terrestrial GNSS interference monitoring from LEO with receiver-reported carrier-to-noise ratio.
- Confirmation of previously-reported interference activity in Syria and Libya, and discovery of previously-unreported ongoing GNSS interference in mainland China.

1.2 Towards Low-Cost Robust PNT

1.2.1 Low-Cost All-Weather Positioning

Development of AGVs has spurred research in lane-keeping assist systems, automated intersection management [53], tight-formation platooning, and cooperative sensing [37, 77], all of which demand accurate (e.g., 50-cm at 95%) ground vehicle positioning in an urban environment. AGVs are currently being operated almost exclusively in areas with dry and sunny climes. Adoption of AGVs in many other parts of the world will demand robustness to hostile weather elements. But the majority of positioning techniques developed thus far depend on lidar or cameras, which perform poorly in low-visibility conditions such as snowy whiteout, dense fog, or heavy rain.

Radio-wave-based sensing techniques such as radar and GNSS remain operable even in extreme weather conditions [173] because their longer-wavelength electromagnetic radiation penetrates snow, fog, and rain. However,

as mentioned earlier, GNSS is either inaccurate or unavailable in deep urban canyons. Additionally, use of automotive radar for localization faces the significant challenges of data sparsity and noise: an automotive radar scan has vastly lower resolution than a camera image or a dense lidar scan, and is subject to high rates of false detection (clutter) and missed detection.

Publicly-available information suggests that no AGV on the road today makes use of radar for all-weather positioning, and no prior literature has demonstrated lane-level-accurate positioning based on automotive radar. In Chap. 4, this dissertation shows that given a prior map of radar reflectors, sub-50-cm accurate all-weather positioning is achievable with low-cost automotive sensors. This work has been published in [109], and makes the following contributions:

- A correlation-maximization-based globally-optimal radar scan registration algorithm applicable to the highly sparse and cluttered data produced by commercially-available low-cost automotive radars.
- A technique for optimally combining sequential radar target estimates so that they are spatially registered with metric consistency. The technique draws on inertial measurements, radar range rate measurements, ground vehicle dynamics constraints, and cm-accurate GNSS measurements, when available.
- A technique for online estimation of the vehicle center of rotation for effective application of ground vehicle dynamics constraints.

- Evaluation of the full positioning pipeline on an urban driving dataset, showing that it maintains 95^{th}-percentile errors below 35 cm in horizontal position and 0.5° in heading during 60 min of GNSS-denied driving.

1.2.2 Crowd-Sourced Mapping & Localization

Mapping the quasi-static driving environment is key to robust positioning and navigation systems, including the radar-based positioning engine mentioned above. These so-called *digital high-definition (HD) maps* for AGVs have many other applications beyond localization: responding to traffic signs and signals, for example, is greatly simplified if the vehicle has prior knowledge of where such signs are located. In short, a map of the surrounding environment enables an AGV to *expect the expected*.

Generating and maintaining these HD maps, however, is a major challenge. Most AGV manufacturers, such as Waymo and General Motors, deploy specialized fleets of mapping vehicles. Generating a map of the environment requires precise knowledge of the vehicle pose (position and orientation) that must be obtained with either an expensive tactical-grade inertial navigation system (INS) or with a high-resolution lidar in a simultaneous localization and mapping (SLAM) framework, or a combination thereof. Furthermore, the map must be updated whenever the quasi-static environment changes, e.g., due to construction. It is time-consuming and impractical to maintain HD maps of entire continents.

A key enabler for large-scale up-to-date maps will be enlisting the help

of the very consumer vehicles that need the maps to build and update them. But connected and automated consumer vehicles will likely be equipped only with low-cost consumer-grade sensor suites. Accordingly, this dissertation explores the accuracy limit of globally-referenced mapping involving collaborating consumer vehicles (Chap. 5). Additionally, it demonstrates a stereo-camera-based digital mapping pipeline called GEOSLAM (globally-referenced electro-optical simultaneous localization and mapping) that achieves the accuracy limit of digital mapping with low-cost automotive sensors (Chap. 6). This work has been published in [113], and makes the following contributions:

- Determination of the asymptotic average error statistics of code-phase GNSS positioning. These statistics govern the accuracy limit of crowd-sourced mapping with code-phase-based GNSS.
- A SLAM pipeline called GEOSLAM that generates a jointly optimized crowd-sourced localization map with mass-market GNSS and visible-light cameras.
- Evaluation of GEOSLAM on a multi-day dataset, showing that sub-50-cm mapping and localization accuracy is achieved after joint optimization over time-separated GNSS measurements.

1.3 Organization

Chaps. 2 and 3 present the contributions in PNT security. The theory of provably secure clock synchronization is detailed in Chap. 2, and a global

survey of GNSS interference is presented in Chap. 3. Chaps. 4, 5, and 6 present the contributions in robust positioning for AGVs. In particular, Chap. 4 develops and demonstrates a prior-map-based all-weather sub-50-cm-accurate localization system, and Chaps. 5 and 6 describe how such a prior map can be crowd-sourced from low-cost consumer vehicles to enable lane-level positioning at scale. Chap. 7 concludes the dissertation.

1.4 List of Publications

1.4.1 Journal Publications

[J1] **Lakshay Narula**, Peter A. Iannucci, and Todd E. Humphreys. All-weather sub-50-cm radar-inertial positioning. *Field Robotics*, 2020. Submitted for review.

[J2] **Lakshay Narula** and Todd E. Humphreys. Requirements for secure clock synchronization. *IEEE Journal of Selected Topics in Signal Processing*, 12(4):749–762, Aug. 2018.

[J3] **Lakshay Narula**, J. Michael Wooten, Matthew J. Murrian, Daniel M. LaChapelle, and Todd E. Humphreys. Accurate collaborative globally-referenced digital mapping with standard GNSS. *Sensors*, 18(8), 2018.

[J4] Matthew J. Murrian, **Lakshay Narula**, Peter A. Iannucci, Scott Budzien, Brady W. O'Hanlon, Steven P. Powell, and Todd E. Humphreys. GNSS interference monitoring from low Earth orbit. *Navigation, Journal of the Institute of Navigation*, 2020. Submitted for review.

[J5] William A. Lies, **Lakshay Narula**, Peter A. Iannucci, and Todd E. Humphreys. Long-range, low SWaP-C FMCW radar. *IEEE Journal of Selected Topics in Signal Processing*, 2020. Submitted for review.

[J6] Todd E. Humphreys, Matthew J. Murrian, and **Lakshay Narula**. Deep urban unaided precise Global Navigation Satellite System vehicle positioning. *IEEE Intelligent Transportation Systems Magazine*, 2020.

1.4.2 Conference Publications

[C1] **Lakshay Narula**, Peter A. Iannucci, and Todd E. Humphreys. Automotive-radar-based 50-cm urban positioning. In *Proceedings of the IEEE/ION PLANSx Meeting*, 2020.

[C2] **Lakshay Narula**, Daniel M. LaChapelle, Matthew J. Murrian, J. Michael Wooten, Todd E. Humphreys, Jean-Baptiste Lacambre, Elliot de Toldi, and Guirec Morvant. TEX-CUP: The University of Texas Challenge for Urban Positioning. In *Proceedings of the IEEE/ION PLANSx Meeting*, 2020.

[C3] **Lakshay Narula**, Matthew J. Murrian, and Todd E. Humphreys. Accuracy limits for globally-referenced digital mapping using standard GNSS. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 3075–3082. IEEE, 2018.

[C4] **Lakshay Narula** and Todd E. Humphreys. Requirements for secure wireless time transfer. In *Proceedings of the IEEE/ION PLANS Meeting*,

Savannah, GA, 2016.

[C5] Matthew J. Murrian, **Lakshay Narula**, and Todd E. Humphreys. Characterizing terrestrial GNSS interference from low Earth orbit. In *Proceedings of the ION GNSS+ Meeting*, Miami, FL, 2019.

[C6] Peter A. Iannucci, **Lakshay Narula**, and Todd E. Humphreys. Cross-modal localization: Using automotive radar for absolute geolocation within a map produced with visible-light imagery. In *Proceedings of the IEEE/ION PLANSx Meeting*, 2020.

[C7] William A. Lies, **Lakshay Narula**, Peter A. Iannucci, and Todd E. Humphreys. Low SWaP-C radar for urban air mobility. In *Proceedings of the IEEE/ION PLANSx Meeting*, 2020.

[C8] Daniel LaChapelle, Todd E. Humphreys, **Lakshay Narula**, Peter A. Iannucci, and Ehsan Moradi-Pari. Automotive collision risk estimation under cooperative sensing. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Barcelona, Spain, 2020.

[C9] Todd E. Humphreys, Matthew J. Murrian, and **Lakshay Narula**. Low-cost precise vehicular positioning in urban environments. In *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pages 456–471, April 2018.

1.4.3 Trade Magazine Publications

[T1] Matthew J. Murrian, **Lakshay Narula**, Todd E. Humphreys, Brady W. O'Hanlon, and Scott Budzien. Characterizing GNSS interference from low-Earth orbit. *Inside GNSS*, 15(1):54–59, 2020.

Part II

Towards Secure PNT

Chapter 2

Requirements for Secure Clock Synchronization

2.1 Abstract

This chapter establishes a fundamental theory of secure clock synchronization. Accurate clock synchronization is the backbone of systems managing power distribution, financial transactions, telecommunication operations, database services, etc. Some clock synchronization (time transfer) systems, such as the GNSS, are based on one-way communication from a master to a slave clock. Others, such as the Network Transport Protocol (NTP), and the IEEE 1588 Precision Time Protocol (PTP), involve two-way communication between the master and slave. This chapter shows that all one-way time transfer protocols are vulnerable to replay attacks that can potentially compromise timing information. A set of conditions for secure two-way clock synchronization is proposed and proved to be necessary and sufficient. It is

This chapter is based on: Lakshay Narula and Todd E. Humphreys. Requirements for secure clock synchronization. *IEEE Journal of Selected Topics in Signal Processing*, 12(4):749–762, Aug. 2018.

shown that IEEE 1588 PTP, although a two-way synchronization protocol, is not compliant with these conditions, and is therefore insecure. Requirements for secure IEEE 1588 PTP are proposed, and a second example protocol is offered to illustrate the range of compliant systems.

2.2 Introduction

Secure clock synchronization is critical to a host of technologies and infrastructure today. The phasor measurement units (PMUs) that enable monitoring and control in power grids need timing information to synchronize measurements across a wide geographical area [120]. Wireless communication networks synchronize their base stations to enable call handoff [93]. Financial networks transfer time across the globe to ensure a common time for pricing and transaction time-stamping [8]. Cloud database services such as Google's Cloud Spanner similarly require precise synchronization between the data centers to maintain consistency [42]. These clock synchronization applications have sub-millisecond accuracy and stringent security requirements.

Clock synchronization is performed either by over-the-wire packet-based communication (NTP, PTP, etc.), or by over-the-air radio signals (GNSS [93], cellular signals, LORAN [138], DCF77 [15], etc.); both wired and wireless clock synchronization are used extensively. Synchronization by GNSS is the method of choice in systems with the most stringent accuracy requirements. Equipped with atomic clocks synchronized to the most accurate time standards available, GNSS satellites can synchronize any number of stations on

Earth to within a few tens of nanoseconds [4]. NTP is usually only accurate to a few milliseconds, but essentially comes for free whenever the host device is connected to a network.

One-way clock synchronization protocols are based on unidirectional communication from the time master station, `A`, to the slave station, `B`. In such protocols, `A` acts as a broadcast station and may send out timing signals either continuously or periodically. The principal drawback of one-way wireless clock synchronization protocols is their vulnerability to delay attacks in which a man-in-the-middle (MITM) adversary nefariously delays or repeats a valid transmission from one station to another. Cryptographic and other measures can improve the security of one-way protocols against delay and other signal- and data-level spoofing attacks [38,123,164], but, as will be shown, such protocols remain fundamentally insecure because of their inability to measure round trip time. They can be secured against unsophisticated attacks, but remain vulnerable to more powerful adversaries.

Two-way clock synchronization protocols involve bi-directional communication between stations `A` and `B`. Such protocols enable measurement of the round trip time of the timing signal, which is shown to be necessary for detecting MITM delay attacks. This measurement, however, is not by itself sufficient for provable security against such attacks.

This chapter establishes a fundamental theory of secure clock synchronization. In contrast to the current literature on timing security [19,97,98,115, 154,158,171], the problem is formalized with definitions, explicit assumptions,

and proofs. The major contributions of this work are as follows:

[T1] One-way synchronization protocols are shown to be insecure against a MITM delay attack. Adversarial delay is shown to be indistinguishable from clock bias, and hence is unobservable without further assumptions.

[T2] A set of necessary conditions for secure two-way clock synchronization is presented and proved. Similar protocol-specific conditions have been previously proposed [9, 98, 158], but have not been generalized to apply to a universal clock synchronization model.

[T3] The proposed necessary conditions, with stricter upper bounds, are shown to be sufficient for secure synchronization in presence of a probabilistic polynomial time (PPT) adversary. Provable security for clock synchronization has not previously been explored in the literature.

[T4] The two-way synchronization scheme of IEEE 1588 PTP is shown to violate a necessary condition for security. This is a known vulnerability of PTP for which a fix has been proposed [158]. Having established a theory for security, this chapter is able to show that the proposed fix is sufficient but is not the minimal necessary modification. A more parsimonious security requirement for PTP is presented that is both necessary and sufficient for secure synchronization.

[T5] A generic construction of a secure two-way clock synchronization protocol is presented to illustrate the general applicability of the proposed necessary and sufficient conditions to a range of underlying protocols.

Wired clock synchronization is inherently more secure than its wireless counterpart because physical access to cables is easier controlled than access to radio channels. This chapter primarily focuses on the more challenging task of clock synchronization over a wireless channel; nonetheless, the attacks and security protocols discussed herein also apply to wireline clock synchronization protocols in the case where the adversary gets access to the channel. For example, if an adversary is able to hijack a boundary clock in a wireline PTP network, then the resulting vulnerabilities are equivalent to that of wireless synchronization where the adversary has open access to the radio channel. In fact, an adversarial boundary clock is even more potent than a wireless adversary since it can completely block the authentic signal from reaching `B`.

The rest of this chapter is organized as follows. Previous works on secure clock synchronization, and their relation to this chapter, are summarized in Section 2.3. Section 2.4 presents a generic model for clock synchronization and shows that all possible one-way synchronization protocols are insecure. Section 2.5 presents the set of security conditions for a wireless clock synchronization protocol, proving these to be necessary by contradiction. Section 2.6 presents a proof of sufficiency for the same set of conditions with stricter upper bounds. A construction of an example secure protocol is presented in Section 2.7, along with the security requirements for IEEE 1588 PTP. Section 2.8 presents a simulation study of a secure clock synchronization model operating over a simplistic channel model. Concluding remarks are made in Section 2.9.

2.3 Related Work

GNSS, NTP, and PTP are the most widely used protocols for clock synchronization. A number of research efforts have been made to assess and improve the security of these protocols. This section reviews some of the notable efforts in the literature.

The GNSS jamming and spoofing threat has been recognized in the literature for more than a decade. A survey of the current state-of-the-art in spoofing and anti-spoofing techniques is presented in [123]. Recent works on GNSS anti-spoofing techniques have specifically focused on the case of timing security. Collaborative multi-receiver [19] and direct time estimation [115] techniques have been proposed for robust GNSS clock synchronization.

The growing popularity of IEEE 1588 PTP for synchronization in critical infrastructure has brought about concerns regarding its security [97, 98, 154, 158, 171]. The threats to IEEE 1588 PTP can broadly be categorized into data-level attacks (such as modification of time messages) and physical layer attacks (such as replay and delay attacks). While cryptographic protocols are able to foil data-level attacks against realistic adversaries, some signal-level attacks, such as the delay attack, remain open vulnerabilities. Unfortunately, their execution is relatively simple. Signal-level attacks, such as the man-in-the-middle attack, have been studied in the recent past. However, these studies only include a brief discussion on countermeasure techniques, and no proof or theoretical guarantee of the efficacy of the countermeasures has been provided.

Ullman et al. [158] propose measuring the propagation delays during initialization of clock synchronization and monitoring the propagation delays during the normal operation of the time synchronization protocol. However, [158] does not prove that such a defense would be sufficient to prevent the delay attacks.

In [98], it is remarked that the clock offset computed between multiple master clocks over a symmetric channel must be zero, and thus, if multiple master clocks are available, they can detect any malicious delay introduced by an adversary. However, this defense does not consider the possibility that the adversary may only delay the packets sent to the slave nodes.

The work presented in [9] is perhaps in closest relation to the current chapter. Annessi et al. upper bound the clock drift between subsequent synchronization signals using a drift model, and perform two-way exchange of timestamps such that the master clock is able to verify the time at the slave. Furthermore, given the maximum clock drift rate and the maximum and minimum propagation delay of the timing signal, they derive an upper bound on the adversarial delay that can go unnoticed. However, with conservative bounds on the maximum clock drift rate and the variation in path delays, the accuracy guarantees derived in [9] may be insufficient for certain applications. Moreover, as will be shown in this chapter, they fail to take account of one the necessary conditions for secure synchronization.

This chapter abstracts the clock synchronization model and assesses its security in a generic setting. It is shown that specialization of the generic

security conditions to the particular protocols assessed in the aforementioned efforts leads to solutions similar or identical to those previously advanced. Thus, establishing the fundamental theory of secure clock synchronization also serves to unify the prior work in the literature.

2.4 System Model

A general system model for clock synchronization is shown in Fig. 2.1. The time seeker station, B, wishes to synchronize its clock to that of the time master station, A. For wireless synchronization applications, stations A and B are assumed to have known locations, $\boldsymbol{x}_{\mathtt{A}}$ and $\boldsymbol{x}_{\mathtt{B}}$, respectively. Due to clock imperfections, the time at station B, $t_{\mathtt{B}}$, continuously drifts with respect to $t_{\mathtt{A}}$, the time at station A. Station B seeks to track the relative drift of its clock by an exchange of signals between A and B. Without loss of generality, this chapter assumes $t_{\mathtt{A}}$ is equivalent to true time (relative to some reference epoch), a close proxy for which is GPS system time.

It is assumed that A and B are able to exchange cryptographic keys securely, if required. This exchange may occur over a public channel via a protocol such as the Diffie-Hellman key exchange [94] or via quantum key exchange techniques [18, 48]. Alternatively, symmetric keys for neighboring stations may be loaded at the time of installation.

Station A sends out a *sync* signal, $s_{\mathtt{A}}$, having distinct features which can be disambiguated from one another by observing a window of the signal containing the feature. The transition in $s_{\mathtt{A}}$ marking the beginning of a data

Table 2.1: Notation used in Chapter 2

$\mathtt{A}$	Time master station
$\mathtt{B}$	Time seeker station
$t_{\mathtt{m}}^{\mathtt{m}_i}$	Transmit time, according to $\mathtt{m}$, of its ith signal feature
$t_{\mathtt{n}}^{\mathtt{m}_i}$	Receipt time, according to $\mathtt{n}$, of the ith signal feature transmitted by $\mathtt{m}$
$\tau_{\mathtt{mn}}^{i}$	Delay, in true time, experienced by the ith feature in propagating from $\mathtt{m}$ to $\mathtt{n}$
$\tau_{\mathtt{mn}_{\mathcal{M}}}^{i}$	Component of $\tau_{\mathtt{mn}}^{i}$ introduced by the man-in-the-middle adversary
$\tau_{\mathtt{mn}_{\mathcal{N}}}^{i}$	Component of $\tau_{\mathtt{mn}}^{i}$ due to natural factors, including processing, transmission, and propagation delay
$\bar{\tau}_{\mathtt{mn}}^{i}$	Modeled or *a priori* estimate of $\tau_{\mathtt{mn}_{\mathcal{N}}}^{i}$
$\tilde{\tau}_{\mathtt{mn}_{\mathcal{N}}}^{i}$	$\tau_{\mathtt{mn}_{\mathcal{N}}}^{i} - \bar{\tau}_{\mathtt{mn}}^{i}$
$\tau_{\mathtt{BB}}$	Delay, in true time, between the receipt of *sync* and transmission of *response* at $\mathtt{B}$
$\bar{\tau}_{\mathtt{BB}}$	Delay, according to $\mathtt{B}$, between the receipt of *sync* and transmission of *response* at $\mathtt{B}$
$\tilde{\tau}_{\mathtt{BB}}$	$\tau_{\mathtt{BB}} - \bar{\tau}_{\mathtt{BB}}$
$\Delta t_{\mathtt{AB}}^{i}$	Clock offset between $\mathtt{A}$ and $\mathtt{B}$ at the time of receipt of the ith feature at $\mathtt{B}$
$\Delta \hat{t}_{\mathtt{AB}}^{i}$	$\mathtt{B}$'s best estimate of $\Delta t_{\mathtt{AB}}^{i}$
$w_{\mathtt{mn}}^{i}$	Measurement noise associated with the measured time-of-arrival of the ith signal feature from $\mathtt{m}$ at $\mathtt{n}$
τ_{RTT}^{ij}	Round trip time, in true time, involving the ith and jth signal features of $\mathtt{A}$ and $\mathtt{B}$, respectively
$\bar{\tau}_{\mathrm{RTT}}^{ij}$	Modeled or *a priori* estimate of τ_{RTT}^{ij}
z_{RTT}^{ij}	A noisy measurement of τ_{RTT}^{ij}

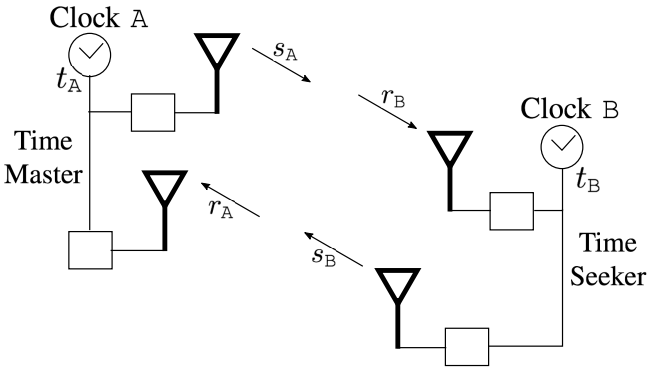


Figure 2.1: Abstract model of a clock synchronization system with a time master station A and a time seeker station B. The antenna outputs are driven by the clock through the receiver and transmitter blocks.

packet is an example of such a signal feature. Furthermore, the system at A is designed such that the kth feature is transmitted at time $t_{\mathtt{A}}^{\mathtt{A}_k}$. B either knows $t_{\mathtt{A}}^{\mathtt{A}_k}$ by prior arrangement, or a digital representation of $t_{\mathtt{A}}^{\mathtt{A}_k}$ is encoded in $s_{\mathtt{A}}$ (e.g., a timestamp). In any case, B knows when the kth feature was sent, according to A's clock. This sets up a bijection

$$S_{\mathtt{A}}^k \rightleftharpoons k \rightleftharpoons t_{\mathtt{A}}^{\mathtt{A}_k} \tag{2.1}$$

where $S_{\mathtt{A}}^k$ represents a window of $s_{\mathtt{A}}$ containing the kth feature.

Station B's received *sync* signal, denoted $r_{\mathtt{B}}$, is a delayed and noisy replica of $s_{\mathtt{A}}$. Let $\tau_{\mathtt{AB}}^k$ denote the delay (in true time) experienced by the kth feature of $s_{\mathtt{A}}$ as it travels from A to B. For line-of-sight (LOS) wireless communication, $\tau_{\mathtt{AB}}^k$ is the sum of the free-space propagation delay over the distance $\|\boldsymbol{x}_{\mathtt{B}} - \boldsymbol{x}_{\mathtt{A}}\|$ and additional delays due to interaction of the timing signal with the intervening channel.

2.4.1 One-Way Clock Synchronization Model

In one-way clock synchronization, the exchange of signals between A and B terminates with reception of the *sync* signal at B. Let $t_{\mathtt{B}}^{\mathtt{A}_k}$ denote the time according to B at which the kth feature of $s_{\mathtt{A}}$ is received at B. The window captured by B containing the kth feature of $s_{\mathtt{A}}$, denoted $R_{\mathtt{B}}^k$, enables B to measure $t_{\mathtt{B}}^{\mathtt{A}_k}$ to within a small error caused by measurement noise. This error, $w_{\mathtt{AB}}^k$, is modeled as zero-mean with variance σ_ϵ^2. The measurement itself, denoted $z_{\mathtt{B}}^k$, is modeled as

$$\begin{aligned} z_{\mathtt{B}}^k &= t_{\mathtt{B}}^{\mathtt{A}_k} + w_{\mathtt{AB}}^k \\ &= t_{\mathtt{A}}^{\mathtt{A}_k} + \tau_{\mathtt{AB}}^k - \Delta t_{\mathtt{AB}}^k + w_{\mathtt{AB}}^k \end{aligned} \tag{2.2}$$

where

$$\Delta t_{\mathtt{AB}}^k \equiv t_{\mathtt{A}}^{\mathtt{A}_k} + \tau_{\mathtt{AB}}^k - t_{\mathtt{B}}^{\mathtt{A}_k} \tag{2.3}$$

is the unknown time offset B wishes to estimate. As the bijection in (2.1) is known to B, B can obtain $t_{\mathtt{A}}^{\mathtt{A}_k}$ for the kth detected feature. If a prior estimate $\bar{\tau}_{\mathtt{AB}}^k$ of the delay $\tau_{\mathtt{AB}}^k$ is available to B, then the desired time offset can be estimated as

$$\Delta\hat{t}_{\mathtt{AB}}^k = t_{\mathtt{A}}^{\mathtt{A}_k} + \bar{\tau}_{\mathtt{AB}}^k - z_{\mathtt{B}}^k \tag{2.4}$$

As a concrete example, consider the case of clock synchronization via GNSS in which B is a GNSS receiver in a known fixed location $\boldsymbol{x}_{\mathtt{B}}$, and A is a GNSS satellite whose location is known to vary with time as $\boldsymbol{x}_{\mathtt{A}}(t_{\mathtt{A}})$. On

detection of the kth feature in a window of captured data, B determines $t_{\mathtt{A}}^{\mathtt{A}_k}$ using (2.1) and also makes the measurement

$$\begin{aligned} z_{\mathtt{B}}^k &= t_{\mathtt{A}}^{\mathtt{A}_k} + \tau_{\mathtt{AB}}^k - \Delta t_{\mathtt{AB}}^k + w_{\mathtt{AB}}^k \\ &= t_{\mathtt{A}}^{\mathtt{A}_k} + \left[\frac{\|\boldsymbol{x}_{\mathtt{B}} - \boldsymbol{x}_{\mathtt{A}}(t_{\mathtt{A}}^{\mathtt{A}_k})\| + D_\rho^k}{c} \right] - \Delta t_{\mathtt{AB}}^k + w_{\mathtt{AB}}^k \end{aligned}$$

where D_ρ^k is the sum of excess ionospheric and neutral-atmospheric delays (in distance units) and c is the speed of light.

The known receiver and satellite positions can be invoked to model the signal's propagation delay as

$$\bar{\tau}_{\mathtt{AB}}^k = \frac{\|\boldsymbol{x}_{\mathtt{B}} - \boldsymbol{x}_{\mathtt{A}}(t_{\mathtt{A}}^{\mathtt{A}_k})\| + \bar{D}_\rho^k}{c}$$

where $\bar{D}_\rho^k$ is a model of the excess delay D_ρ^k at the time of receipt of the kth feature at B. The modeled excess delay is based on atmospheric models possibly refined by dual-frequency measurements [96]. An estimate of the time offset, $\Delta\hat{t}_{\mathtt{AB}}^k$, can then be made using $t_{\mathtt{A}}^{\mathtt{A}_k}$, $z_{\mathtt{B}}^k$, and $\bar{\tau}_{\mathtt{AB}}^k$ in (2.4).

It must be noted that, for one-way clock synchronization, any errors in the estimate of the distance between A and B, and in the estimate of the excess channel delay, will appear as an error in the estimate of the time offset.

2.4.2 Two-Way Clock Synchronization Model

As discussed above, if an estimate of $\bar{\tau}_{\mathtt{AB}}^k$ is available, then clock synchronization is complete after B receives the *sync* signal $r_{\mathtt{B}}$. The *response* signal from B in a two-way protocol is typically used to either determine, or refine,

the estimate of $\bar{\tau}^k_{\mathtt{AB}}$ with a measurement of the round trip time (RTT). The ability to measure RTT obviates the requirement that $\|\boldsymbol{x}_{\mathtt{B}} - \boldsymbol{x}_{\mathtt{A}}\|$ be known *a priori*. In IEEE 1588 PTP, for example, RTT is measured to initially obtain, and periodically refine, the value of $\bar{\tau}^k_{\mathtt{AB}}$ used in deriving $\Delta\hat{t}^k_{\mathtt{AB}}$ from (2.4).

In the system model considered in this chapter, station B transmits a *response* $s_{\mathtt{B}}$ that is designed such that (1) there is a one-to-one mapping $l(k)$ between the lth feature in $s_{\mathtt{B}}$ and the kth feature in $s_{\mathtt{A}}$, and (2) the lth feature's index can be inferred by observation of a window containing it. Symbolically, if $S^l_{\mathtt{B}}$ is a window of $s_{\mathtt{B}}$ containing the lth feature of the *response* signal, then

$$S^l_{\mathtt{B}} \rightleftharpoons l(k) \rightleftharpoons k \tag{2.5}$$

On receipt of the kth feature in $s_{\mathtt{A}}$, at time $t^{\mathtt{A}_k}_{\mathtt{B}}$ by B's clock, but at $z^k_{\mathtt{B}}$ as measured by B, B transmits the lth feature in $s_{\mathtt{B}}$ after a short delay, $\tau_{\mathtt{BB}}$ (in true time), hereon referred to as the *layover time*.

The layover time is introduced as a practical consideration. On receipt of A's kth feature, B is physically unable to transmit its own lth feature with zero delay. Thus, B is allowed to specify a short layover time, $\bar{\tau}_{\mathtt{BB}}$, after which it intends to launch its lth feature. It is important to note that the actual layover time, $\tau_{\mathtt{BB}}$, will not be the same as the intended layover time due to (1) non-zero measurement noise $w^k_{\mathtt{AB}}$ and (2) non-zero frequency offset of the clock at B with respect to true time. However, if the layover time is sufficiently short and the measurement noise is benign, the difference $\bar{\tau}_{\mathtt{BB}} - \tau_{\mathtt{BB}}$ can be made negligible compared to the time synchronization requirement, with the actual

value depending on the quality of B's clock.

Station A receives the *response* signal as a delayed and noisy replica of $s_{\mathtt{B}}$, denoted $r_{\mathtt{A}}$. The delay experienced by the lth feature as it travels from B to A, in true time, is denoted $\tau^l_{\mathtt{BA}}$. Station A captures a window $R^l_{\mathtt{A}}$ of $r_{\mathtt{A}}$ that enables A to identify the lth feature in $s_{\mathtt{B}}$ according to (2.5), and to infer that the received feature is in response to the kth feature transmitted by A. Furthermore, A makes a noise-corrupted measurement $z^l_{\mathtt{A}}$ of the time-of-arrival of the lth feature in $s_{\mathtt{B}}$, according to A's clock. The noise, denoted $w^l_{\mathtt{BA}}$, is again modeled as zero-mean with variance σ^2_ϵ. The full measurement model is given by

$$\begin{aligned} z^l_{\mathtt{A}} &= t_{\mathtt{A}}^{\mathtt{B}_l} + w^l_{\mathtt{BA}} \\ &= t_{\mathtt{A}}^{\mathtt{A}_k} + \tau^k_{\mathtt{AB}} + \tau_{\mathtt{BB}} + \tau^l_{\mathtt{BA}} + w^l_{\mathtt{BA}} \end{aligned}$$

Since $t_{\mathtt{A}}^{\mathtt{A}_k}$ is exactly known at A, a direct noisy measurement of the round trip time $\tau^k_{\mathtt{AB}} + \tau_{\mathtt{BB}} + \tau^l_{\mathtt{BA}}$ can be made as

$$z^{kl}_{\mathrm{RTT}} \equiv z^l_{\mathtt{A}} - t_{\mathtt{A}}^{\mathtt{A}_k} \tag{2.6}$$

Note that the noise $w^l_{\mathtt{BA}}$ and $w^k_{\mathtt{AB}}$ in z^{kl}_{RTT} is embedded within $z^l_{\mathtt{A}}$ and $\tau_{\mathtt{BB}}$, respectively. Under the assumption of symmetric delays, i.e., $\tau^k_{\mathtt{AB}} = \tau^l_{\mathtt{BA}}$, and with knowledge of $\bar{\tau}_{\mathtt{BB}}$, the measured RTT in (2.6) can be exploited to improve the modeled propagation delay for future exchanges between A and B:

$$\bar{\tau}^m_{\mathtt{AB}} = \bar{\tau}^n_{\mathtt{BA}} = \frac{z^{kl}_{\mathrm{RTT}} - \bar{\tau}_{\mathtt{BB}}}{2}$$

where $m > k$ and $n > l$.

The two-way exchange of *sync* and *response* messages is summarized visually in Fig. 2.2.

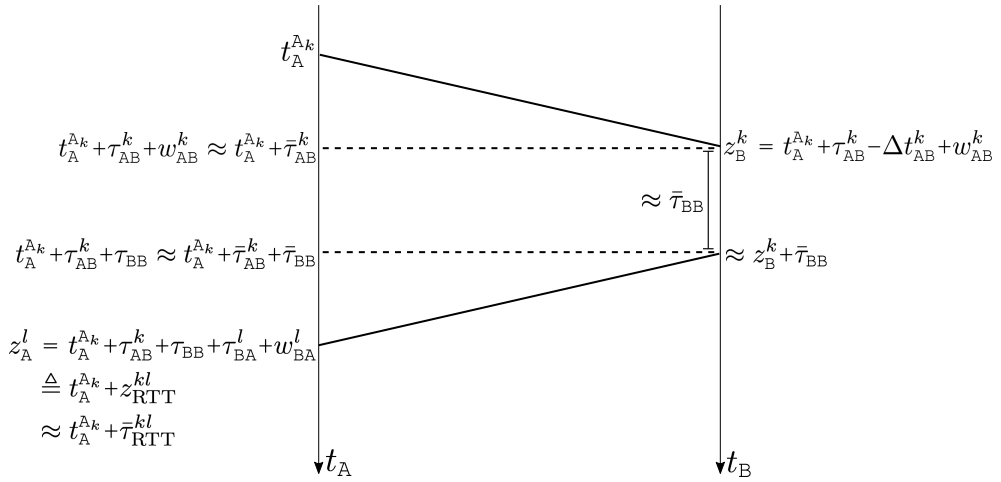


Figure 2.2: Two-way exchange of *sync* and *response* messages between `A` and `B` in the absence of a man-in-the-middle adversary.

Since RTT will play a central role in the discussion on secure synchronization later on, various definitions and assumptions concerning RTT are stated here for clarity:

- RTT for the kth feature in $s_{\mathtt{A}}$ and the corresponding lth feature in $s_{\mathtt{B}}$ is defined as
$$\tau_{\mathrm{RTT}}^{kl} \equiv \tau_{\mathtt{AB}}^{k} + \tau_{\mathtt{BB}} + \tau_{\mathtt{BA}}^{l}$$
- Measured RTT includes, in addition to RTT, measurement noise at `A`; it is modeled as
$$z_{\mathrm{RTT}}^{kl} = \tau_{\mathtt{AB}}^{k} + \tau_{\mathtt{BB}} + \tau_{\mathtt{BA}}^{l} + w_{\mathtt{BA}}^{l}$$

- Modeled RTT, also called the prior estimate of RTT, is defined as

$$\bar{\tau}^{kl}_{\mathrm{RTT}} \equiv \bar{\tau}^{k}_{\mathtt{AB}} + \bar{\tau}_{\mathtt{BB}} + \bar{\tau}^{l}_{\mathtt{BA}} \tag{2.7}$$

  For example, in the case of wireless clock synchronization with LOS electromagnetic signals, a prior estimate of RTT is based on the distance between A and B and on models of channel delays in excess of free-space propagation between these.

- The modeled RTT, $\bar{\tau}^{kl}_{\mathrm{RTT}}$, can be refined with measurements of RTT in a two-way protocol. Alternatively, as will be discussed later, if an accurate modeled RTT is available, it and the measured RTT can be used to detect delay attacks.

- Unambiguous measurement of RTT requires that there exist a one-to-one mapping between the signal features in $s_{\mathtt{A}}$ and $s_{\mathtt{B}}$, as mathematically represented in (2.5). On detection of the lth feature in $s_{\mathtt{B}}$, A must be able to deduce that this feature was transmitted approximately $\bar{\tau}_{\mathtt{BB}}$ after B received the kth feature in $s_{\mathtt{A}}$. This requirement is appropriately a part of the RTT definition since it enables A to unambiguously measure RTT.

2.4.3 Attack Model

The attack model in this chapter considers a MITM adversary $\mathcal{M}$. The available computational resources allow $\mathcal{M}$ to execute probabilistic polynomial time (PPT) algorithms. $\mathcal{M}$ can receive, detect, and replay signals from A and B with arbitrarily precise directional antennas. Additionally, $\mathcal{M}$ has precise

knowledge of $\boldsymbol{x}_{\mathtt{A}}$ and $\boldsymbol{x}_{\mathtt{B}}$, and can take up any position around or between the two stations. It has unrestricted access to the signals that A and B exchange over the air, and has complete knowledge of their synchronization protocol save for the cryptographic keys.

Let L denote the alert limit, defined as the error in time synchronization not to be exceeded without issuing an alert.

Definition 1. *Clock synchronization is defined to be compromised if* $|\Delta t_{\mathtt{AB}} - \Delta\hat{t}_{\mathtt{AB}}| \geq L$.

Note that, in the absence of an adversary, clock synchronization is not compromised so long as

$$|\tau^k_{\mathtt{AB}} - \bar{\tau}^k_{\mathtt{AB}} + w^k_{\mathtt{AB}}| < L$$

However, in the presence of a MITM adversary, the *sync* signal is delayed or advanced such that

$$\tau^k_{\mathtt{AB}} = \tau^k_{\mathtt{AB}_{\mathcal{N}}} + \tau^k_{\mathtt{AB}_{\mathcal{M}}} \tag{2.8}$$

where $\tau^k_{\mathtt{AB}_{\mathcal{N}}} > 0$ is the natural or physical delay (equal to $\tau^k_{\mathtt{AB}}$ in the absence of an adversary) and $\tau^k_{\mathtt{AB}_{\mathcal{M}}} \geq 0$ is the adversarial delay. In this case, if

$$|\tau^k_{\mathtt{AB}} - \bar{\tau}^k_{\mathtt{AB}} + w^k_{\mathtt{AB}}| = |\tau^k_{\mathtt{AB}_{\mathcal{N}}} - \bar{\tau}^k_{\mathtt{AB}} + \tau^k_{\mathtt{AB}_{\mathcal{M}}} + w^k_{\mathtt{AB}}| \geq L \tag{2.9}$$

then clock synchronization is compromised.

2.4.4 Vulnerability of One-Way Clock Synchronization

One-way clock synchronization is fundamentally vulnerable to a delay attack because it provides no mechanism to measure RTT. The adversary

$\mathcal{M}$ can compromise any one-way wireless clock synchronization protocol by retransmitting the authentic *sync* signal from A such that the retransmitted signal, $s_{\mathcal{M}}$, overpowers or otherwise supersedes the authentic signal $s_{\mathtt{A}}$. In the absence of additional assumptions beyond those underpinning the one-way protocol described earlier, $\mathcal{M}$ can introduce an arbitrary delay $\tau^k_{\mathtt{AB}_{\mathcal{M}}}$ in its retransmission, thereby compromising the synchronization process.

Note that whereas counterfeit signal attacks can be prevented by authentication and cryptographic methods [166], these techniques do not prevent delay attacks because the delayed or repeated signal has the same cryptographic characteristics as that of the genuine signal, the only difference being that it is received with a (possibly small) additional delay.

The delay introduced by $\mathcal{M}$ is added to the natural delay, $\tau^k_{\mathtt{AB}_{\mathcal{N}}}$, of the signal between A and B. As a result, an error of $\approx \tau^k_{\mathtt{AB}_{\mathcal{M}}}$ is introduced in the estimated time offset at B. From (2.4), it follows that

$$\begin{aligned}
\Delta\hat{t}^k_{\mathtt{AB}} &= t^{\mathtt{A}_k}_{\mathtt{A}} + \bar{\tau}^k_{\mathtt{AB}} - z^k_{\mathtt{B}} \\
&= t^{\mathtt{A}_k}_{\mathtt{A}} + \bar{\tau}^k_{\mathtt{AB}} - (t^{\mathtt{A}_k}_{\mathtt{A}} + \tau^k_{\mathtt{AB}} - \Delta t^k_{\mathtt{AB}} + w^k_{\mathtt{AB}}) \\
&= (\bar{\tau}^k_{\mathtt{AB}} - \tau^k_{\mathtt{AB}_{\mathcal{N}}}) - \tau^k_{\mathtt{AB}_{\mathcal{M}}} + \Delta t^k_{\mathtt{AB}} - w^k_{\mathtt{AB}} \\
&\approx \Delta t^k_{\mathtt{AB}} - \tau^k_{\mathtt{AB}_{\mathcal{M}}}
\end{aligned} \tag{2.10}$$

where it is assumed that the error due to inaccurately modeled delay is negligible and that $\sigma_\epsilon \ll \tau^k_{\mathtt{AB}_{\mathcal{M}}}$. In the absence of an RTT measurement, and without further assumptions on the nature of the protocol or the clock drift model

considered, the adversarial delay $\tau_{\mathtt{AB}_{\mathcal{M}}}$ is indistinguishable from a clock offset of the same magnitude.

To be sure, measures can be taken to make a MITM delay attack harder to execute without detection. But, importantly, these measures cannot guarantee that the synchronization will remain uncompromised. Various measures proposed in the literature, and their shortcomings, are discussed below.

Received Signal Strength Monitoring The adversary $\mathcal{M}$ might attempt to overpower the authentic signal in order to spoof the *sync* message, leading to an increase in the total signal power received at B. Station B could monitor the received signal strength (RSS) to detect such an attack [2]. However, a potent adversary could transmit, in addition to its delayed signal, an amplitude-matched, phase-inverted nulling signal that annihilates the authentic *sync* signal $s_{\mathtt{A}}$ as received at B, thus preventing an unusual increase in received power at B. If $\mathcal{M}$ is positioned along the straight-line path between A and B, nulling of $s_{\mathtt{A}}$ can be effected without prior knowledge of $s_{\mathtt{A}}$. A laboratory demonstration of such nulling is reported in [65].

Selective Rejection of False Signal If B receives both the authentic and false (delayed) *sync* signals, it may be able to apply angle-of-arrival or signal processing techniques to selectively reject the delayed signal [25, 95, 123, 164]. However, discrimination based on angle-of-arrival fails if $\mathcal{M}$ is positioned along the line from A to B, and, as conceded in [164], signal-processing-based tech-

niques for selective rejection of false signals can be thwarted by an adversary transmitting an additional nulling signal, as described above.

Collaborative Verification Multiple time seekers may attempt to synchronize to the same time master. In this scenario, the time seekers can potentially detect malicious activity by cross-checking the received signals [19]. In the simplest implementation, all time seekers can collaborate to verify that they are synchronized amongst each other. In case of an uncoordinated attack against a subset of time seekers, this verification would expose the attack since the time offset computed at the attacked subset would be different from that computed at the other stations. In principle, however, it is possible for an adversary to execute a coordinated attack against all the time seekers, thus concealing its presence.

2.5 Necessary Conditions for Secure Synchronization

This section presents a set of conditions for secure two-way clock synchronization and proves these to be necessary by contradiction. In other words, it is shown that if a two-way clock synchronization protocol does not satisfy any one of these proposed conditions, there exists an attack that can compromise clock synchronization without detection.

It is important to note that the ability to measure RTT in a two-way protocol is necessary, but not sufficient, for provably secure synchronization. As an example, IEEE 1588 PTP is a two-way protocol that has been proposed

as an alternative to GNSS for sub-microsecond clock synchronization in critical infrastructure such as the PMU network. But, despite the bi-directional exchange between stations, and hence the ability to measure RTT, recent work has shown that PTP is vulnerable to delay attacks in which a MITM introduces asymmetric delay between A and B. Asymmetric delay breaks the assumption that $\tau^k_{\mathtt{AB}} = \tau^k_{\mathtt{BA}}$ and leads to an erroneous prior for $\bar{\tau}_{\mathtt{AB}}$ and $\bar{\tau}_{\mathtt{BA}}$ for future exchanges. This vulnerability is documented in both the literature [9, 98, 158] and the IEEE 1588-2008 standard. Thus, a secure two-way clock synchronization protocol must satisfy additional security requirements beyond the ability to measure RTT.

The conditions introduced below are not tied to any specific protocol, unlike some measures proposed in the current literature [19, 97, 98, 115, 154, 158, 171]. They are generally applicable to any two-way protocol (e.g., PTP) for which the foregoing two-way synchronization model applies.

Assuming the time master A initiates the two-way communication, the necessary conditions for secure clock synchronization are as follows:

[T1] Both A and B must transmit unpredictable waveforms to prevent the adversary $\mathcal{M}$ from generating counterfeit signals that pass authentication. In practice, this implies the use of a cryptographic construct such as a message authentication code (MAC) or a digital signature.

[T2] The propagation time of the signal must be irreducible to within the alert limit L along both signal paths. For wireless clock synchronization,

this condition implies synchronization via LOS electromagnetic signals as $L \to 0$.

[T3] The RTT between A and B must be known to A and measurable by A to within the alert limit L. The RTT must include the delays internal to both A and B, in addition to the propagation delay. Station A must know of any intentional delay introduced by B, such as the layover time $\tau_{\mathtt{BB}}$ introduced earlier.

2.5.1 Proof of Necessity of Conditions

Stations A and B must transmit unpredictable signals To prove this condition is necessary, two scenarios are considered: *a)* station A transmits a signal waveform $s_{\mathtt{A}}$ that is predictable, and, *b)* station B transmits a signal waveform $s_{\mathtt{B}}$ that is predictable.

$s_{\mathtt{A}}$ is predictable $\mathcal{M}$ can compromise synchronization without detection as follows:

i) $\mathcal{M}$ takes up a position between A and B along the line joining the antennas at the two stations.

ii) $\mathcal{M}$ initially transmits a replica of $s_{\mathtt{A}}$ such that B receives identical signals from both A and $\mathcal{M}$. Subsequently, $\mathcal{M}$ increases its signal power or otherwise supersedes $s_{\mathtt{A}}$ (e.g., via signal nulling, as discussed earlier) such that B tracks $s_{\mathcal{M}}$, the signal transmitted by $\mathcal{M}$. (Hereafter, whenever

signals from $\mathcal{M}$ compete with those from A or B, it will be assumed that those from $\mathcal{M}$ exert control.)

iii) Exploiting the predictability of $s_{\mathtt{A}}$, $\mathcal{M}$ advances its replica $s_{\mathcal{M}}$ with respect to $s_{\mathtt{A}}$ by $|\tau^k_{\mathtt{AB}_{\mathcal{M}}}|$, where $\tau^k_{\mathtt{AB}_{\mathcal{M}}} < 0$. B tracks the advanced signal, resulting in an error of $\tau^k_{\mathtt{AB}_{\mathcal{M}}}$ in the computed $\Delta\hat{t}^k_{\mathtt{AB}}$ as shown in (2.10).

iv) B transmits the unpredictable *response* $s_{\mathtt{B}}$ compliant with the prearranged layover time $\bar{\tau}_{\mathtt{BB}}$. $\mathcal{M}$ intercepts this signal from B, and replays it to A with a delay of $\tau^l_{\mathtt{BA}_{\mathcal{M}}} = -\tau^k_{\mathtt{AB}_{\mathcal{M}}} > 0$, causing A to track the delayed signal. As a result, the RTT is $\tau^k_{\mathtt{AB}} + \tau_{\mathtt{BB}} + \tau^l_{\mathtt{BA}}$ as A expects. In summary:

$$\begin{aligned} \tau^k_{\mathtt{AB}} &= \tau^k_{\mathtt{AB}_{\mathcal{N}}} + \tau^k_{\mathtt{AB}_{\mathcal{M}}} \\ \tau^l_{\mathtt{BA}} &= \tau^l_{\mathtt{BA}_{\mathcal{N}}} + \tau^l_{\mathtt{BA}_{\mathcal{M}}} = \tau^l_{\mathtt{BA}_{\mathcal{N}}} - \tau^k_{\mathtt{AB}_{\mathcal{M}}} \\ \Rightarrow \tau^k_{\mathtt{AB}} + \tau^l_{\mathtt{BA}} &= \tau^k_{\mathtt{AB}_{\mathcal{N}}} + \tau^l_{\mathtt{BA}_{\mathcal{N}}} \end{aligned}$$

Thus, $\mathcal{M}$ undoes the effect of its *sync* advance, preventing A from detecting the attack.

$s_{\mathtt{B}}$ is predictable $\mathcal{M}$ can compromise synchronization without detection by replicating B's behavior:

i) $\mathcal{M}$ takes up a position between A and B along the line joining the antennas at the two stations.

ii) $\mathcal{M}$ receives the *sync* signal and generates a valid *response* with a delay

$$\bar{\tau}_{\mathtt{BB}} + \frac{\|\boldsymbol{x}_{\mathcal{M}} - \boldsymbol{x}_{\mathtt{B}}\|}{\|\boldsymbol{x}_{\mathtt{A}} - \boldsymbol{x}_{\mathtt{B}}\|}\left(\bar{\tau}^k_{\mathtt{AB}} + \bar{\tau}^l_{\mathtt{BA}}\right) \tag{2.11}$$

such that the RTT is $\bar{\tau}^k_{\mathtt{AB}} + \bar{\tau}_{\mathtt{BB}} + \bar{\tau}^l_{\mathtt{BA}}$, as A expects.

iii) $\mathcal{M}$ records the unpredictable signal from A and replays it to B with an arbitrary delay $\tau^k_{\mathtt{AB}_{\mathcal{M}}} > 0$. This results in an error of approximately $\tau^k_{\mathtt{AB}_{\mathcal{M}}}$ in the computed $\Delta\hat{t}^k_{\mathtt{AB}}$ at B, as shown in (2.10).

Propagation time must be irreducible to within L If there exists a channel that reduces the propagation time from A to B or from B to A by more than L as compared to the channel used by A and B, then $\mathcal{M}$ can compromise synchronization without detection. The following attack assumes the propagation time from A to B is reducible by more than L; a similar attack exploits the situation in which the propagation time from B to A is reducible by more than L.

i) $\mathcal{M}$ records the *sync* signal $s_{\mathtt{A}}$ going from A to B.

ii) $\mathcal{M}$ makes the recorded signal reach B advanced by $|\tau^k_{\mathtt{AB}_{\mathcal{M}}}|$ compared to $s_{\mathtt{A}}$, where $\tau^k_{\mathtt{AB}_{\mathcal{M}}} < -L$. An error of $\tau^k_{\mathtt{AB}_{\mathcal{M}}}$ is introduced in the time offset value computed at B as shown in (2.10).

iii) $\mathcal{M}$ records the *response* signal $s_{\mathtt{B}}$, which has the expected prearranged layover time $\tau_{\mathtt{BB}} \approx \bar{\tau}_{\mathtt{BB}}$. $\mathcal{M}$ replays this signal to A with a delay of $\tau^l_{\mathtt{BA}_{\mathcal{M}}} = -\tau^k_{\mathtt{AB}_{\mathcal{M}}}$ such that the RTT is consistent with what A expects.

RTT known to and measurable by A to within L Synchronization can be compromised without detection if $|z^{kl}_{\mathrm{RTT}} - \bar{\tau}^{kl}_{\mathrm{RTT}}| > L$ with non-negligible

probability even in the absence of an adversary. This condition can be met if *a)* the prior estimates $\bar{\tau}^k_{\mathtt{AB}}$, $\bar{\tau}^l_{\mathtt{BA}}$, or $\bar{\tau}_{\mathtt{BB}}$ are not accurate to the corresponding true values to within L, or *b)* the magnitude of the measurement error sum $|w^k_{\mathtt{AB}} + w^l_{\mathtt{BA}}|$ is larger than L. Note that the condition $|w^k_{\mathtt{AB}}| > L$ compromises synchronization even absent an adversary. An adversary $\mathcal{M}$ can exploit the condition $|z^{kl}_{\mathrm{RTT}} - \bar{\tau}^{kl}_{\mathrm{RTT}}| > L$ as follows:

i) $\mathcal{M}$ initially transmits a replica of $s_{\mathtt{A}}$ such that B receives identical signals from both A and $\mathcal{M}$. Subsequently, $\mathcal{M}$ introduces a delay $\tau^k_{\mathtt{AB}_{\mathcal{M}}} > 0$ in the replayed signal $s_{\mathcal{M}}$. As assumed earlier, $s_{\mathcal{M}}$ exerts control and introduces an error of approximately $\tau^k_{\mathtt{AB}_{\mathcal{M}}}$ in the computed $\Delta\hat{t}^k_{\mathtt{AB}}$ at B, as shown in (2.10).

ii) Station B transmits the *response* signal with the prearranged layover time $\tau_{\mathtt{BB}} \approx \bar{\tau}_{\mathtt{BB}}$ with respect to the delayed signal.

iii) In the received signal $r_{\mathtt{A}}$, A identifies the expected feature $l(k)$. The RTT, if measurable, includes the delay $\tau^k_{\mathtt{AB}_{\mathcal{M}}}$ introduced by $\mathcal{M}$.

iv) However, A is unable to definitively declare an attack, since the errors in the modeled RTT and/or the measurement of RTT are possibly larger than L. In other words, it is not possible to claim that $|z^{kl}_{\mathrm{RTT}} - \bar{\tau}^{kl}_{\mathrm{RTT}}| > L$ only in the presence of adversarial delay.

2.6 Proof of Sufficiency

This section presents a sufficiency proof for the set of security conditions proposed in the previous section. A sufficiency proof guarantees secure synchronization under the considered system and attack models. This chapter draws inspiration from the literature on modern cryptography and formalizes the problem of secure clock synchronization with explicit definitions, assumptions, and proofs.

2.6.1 Assumptions

This proof assumes that the system under consideration strictly complies with the set of necessary security conditions. Specifically,

[T1] Both `A` and `B` use an authenticated encryption scheme to generate unpredictable and verifiably authentic signals in the presence of a probabilistic polynomial time (PPT) adversary.

[T2] The difference between the RTT along the communication channel between `A` and `B` and the shortest possible RTT is negligible as compared to L.

[T3] The difference between the modeled delays $\bar{\tau}^k_{\mathtt{AB}}$ and $\bar{\tau}^l_{\mathtt{BA}}$ and the true delays $\tau^k_{\mathtt{AB}}$ and $\tau^l_{\mathtt{BA}}$, respectively, is negligible as compared to L.

$$|\bar{\tau}^k_{\mathtt{AB}} - \tau^k_{\mathtt{AB}_{\mathcal{N}}}| \ll L \tag{2.12}$$

and

$$|\bar{\tau}^l_{\mathtt{BA}} - \tau^l_{\mathtt{BA}_{\mathcal{N}}}| \ll L \tag{2.13}$$

Furthermore, A and B agree upon a fixed layover time $\bar{\tau}_{\mathtt{BB}}$, and the difference between this and the true layover time is negligible: $|\tau_{\mathtt{BB}} - \bar{\tau}_{\mathtt{BB}}| \ll L$.

[T4] The standard deviation of the noise corrupting the measurements $t_{\mathtt{B}}^{\mathtt{A}_k}$ and $t_{\mathtt{A}}^{\mathtt{B}_l}$ is negligible compared to the alert limit:

$$\sigma_\epsilon \ll L \tag{2.14}$$

Notice that the above assumptions are the same as the necessary conditions in Section 2.5, but with stricter upper bounds on the conditions.

If symmetric keys are exchanged prior to synchronization, then private-key cryptographic schemes such as Encrypt-then-MAC [17] can be used for authenticated encryption. Alternatively, if the keys must be exchanged over a public channel, then digital signatures [59] can be used to authenticate the encrypted messages. Cryptographic authentication schemes like MAC and digital signatures generate a tag associated with a message. Qualitatively, a MAC or digital signature scheme is secure if a PPT adversary, even when given access to multiple valid message-tag pairs of its own choice (as many as possible in polynomial time), cannot generate a valid tag for a new message with non-negligible probability. Irrespective of the cryptographic scheme used, this proof assumes that the probability of $\mathcal{M}$ generating a new valid *sync* or *response* signal is a negligible function of the key length n:

$$\mathbb{P}[\mathsf{Valid}] < \mathsf{negl}(n) \tag{2.15}$$

To detect an attack before the synchronization error exceeds L, A must select a threshold lower than L beyond which an attack is declared. Consider the

modeled RTT, $\bar{\tau}^{kl}_{\mathrm{RTT}}$, as defined in (2.7), and the measurement z^{kl}_{RTT} as defined in (2.6). A threshold less than L, say $L - \delta$ with $0 < \delta < L$, is set by station A such that if $|z^{kl}_{\mathrm{RTT}} - \bar{\tau}^{kl}_{\mathrm{RTT}}| > L - \delta$, then an attack is declared.

2.6.2 Definitions

Definition 2. *A PPT adversary $\mathcal{M}$ succeeds if clock synchronization is compromised (Definition 1) and*

$$|z^{kl}_{\mathrm{RTT}} - \bar{\tau}^{kl}_{\mathrm{RTT}}| \leq L - \delta$$

Definition 3. *Faster-than-light (superluminal) propagation is defined to be hard if $\mathcal{M}$ cannot propagate a signal at a speed higher than the speed of light with non-negligible probability. Under hardness of superluminal propagation*

$$\mathbb{P}[\mathsf{Superluminal}] \approx 0$$

Definition 4. *A clock synchronization protocol is defined to be secure if, under the hardness of superluminal propagation assumption,*

$$\mathbb{P}[\mathsf{Success}] < \mathsf{negl}(n)$$

where Success *for $\mathcal{M}$ is defined in Definition 2.*

2.6.3 Proof

In the presence of an adversary $\mathcal{M}$, the measurement z^{kl}_{RTT} is modeled as

$$z^{kl}_{\mathrm{RTT}} = \tau^{k}_{\mathtt{AB}_{\mathcal{N}}} + \tau^{k}_{\mathtt{AB}_{\mathcal{M}}} + \tau^{l}_{\mathtt{BA}_{\mathcal{N}}} + \tau^{l}_{\mathtt{BA}_{\mathcal{M}}} + \tau_{\mathtt{BB}} + w^{l}_{\mathtt{BA}} \tag{2.16}$$

Let $\tilde{\tau}^{k}_{\mathtt{AB}_{\mathcal{N}}}$ and $\tilde{\tau}^{l}_{\mathtt{BA}_{\mathcal{N}}}$ denote the error in the modeled signal delay due to natural/physical phenomenon. Also, let $\tilde{\tau}_{\mathtt{B}}$ be the difference between the intended layover time $\bar{\tau}_{\mathtt{BB}}$ and the actual layover time $\tau_{\mathtt{BB}}$. Note that these might be positive or negative.

$$\tilde{\tau}^{k}_{\mathtt{AB}_{\mathcal{N}}} = \tau^{k}_{\mathtt{AB}_{\mathcal{N}}} - \bar{\tau}^{k}_{\mathtt{AB}} \tag{2.17}$$

$$\tilde{\tau}^{l}_{\mathtt{BA}_{\mathcal{N}}} = \tau^{l}_{\mathtt{BA}_{\mathcal{N}}} - \bar{\tau}^{l}_{\mathtt{BA}} \tag{2.18}$$

$$\tilde{\tau}_{\mathtt{BB}} = \tau_{\mathtt{BB}} - \bar{\tau}_{\mathtt{BB}} \tag{2.19}$$

From (2.7), (2.16), (2.17), (2.18), and (2.19) it follows that

$$z^{kl}_{\mathrm{RTT}} = \bar{\tau}^{kl}_{\mathrm{RTT}} + \tilde{\tau}^{k}_{\mathtt{AB}_{\mathcal{N}}} + \tau^{k}_{\mathtt{AB}_{\mathcal{M}}} + \tilde{\tau}^{l}_{\mathtt{BA}_{\mathcal{N}}} + \tau^{l}_{\mathtt{BA}_{\mathcal{M}}} + \tilde{\tau}_{\mathtt{BB}} + w^{l}_{\mathtt{BA}}$$

Following the assumptions in (2.12) and (2.13), the residual delays are negligible in comparison to L:

$$|\tilde{\tau}^{k}_{\mathtt{AB}_{\mathcal{N}}}| \ll L \tag{2.20}$$

$$|\tilde{\tau}^{l}_{\mathtt{BA}_{\mathcal{N}}}| \ll L \tag{2.21}$$

This assumption is reasonable since otherwise the system could not confidently meet the accuracy requirements even in the absence of an adversary. Also, if

$\bar{\tau}_{\mathtt{BB}}$ is a short time interval and the measurement noise σ_ϵ is benign, it is reasonable to assume that

$$|\tilde{\tau}_{\mathtt{BB}}| \ll L \tag{2.22}$$

Note that $\mathcal{M}$ can advance the signal by (a) forging a valid message/tag pair, or (b) propagating the signal faster-than-light. The assumptions of secure MAC and hardness of superluminal propagation enforce that

$$\begin{aligned}\mathbb{P}[\tau^k_{\mathtt{AB}_{\mathcal{M}}} < 0] &< \mathbb{P}[\mathsf{Valid}] + \mathbb{P}[\mathsf{Superluminal}] \\ &\approx \mathsf{negl}(n)\end{aligned}$$

In order to stay undetected, the adversary must ensure

$$\begin{aligned}L - \delta &\geq |z^{kl}_{\mathrm{RTT}} - \bar{\tau}^{kl}_{\mathrm{RTT}}| \\ &= |\tilde{\tau}^k_{\mathtt{AB}_{\mathcal{N}}} + \tau^k_{\mathtt{AB}_{\mathcal{M}}} + \tilde{\tau}^l_{\mathtt{BA}_{\mathcal{N}}} + \tau^l_{\mathtt{BA}_{\mathcal{M}}} + \tilde{\tau}_{\mathtt{BB}} + w^l_{\mathtt{BA}}|\end{aligned} \tag{2.23}$$

At the same time, in order to compromise time transfer, from (2.9), $\mathcal{M}$ must ensure

$$\begin{aligned}L &\leq |\tilde{\tau}^k_{\mathtt{AB}_{\mathcal{N}}} + \tau^k_{\mathtt{AB}_{\mathcal{M}}} + w^k_{\mathtt{AB}}| \\ &\leq |\tilde{\tau}^k_{\mathtt{AB}_{\mathcal{N}}} + w^k_{\mathtt{AB}}| + |\tau^k_{\mathtt{AB}_{\mathcal{M}}}| \\ \Rightarrow |\tau^k_{\mathtt{AB}_{\mathcal{M}}}| &\geq L - |\tilde{\tau}^k_{\mathtt{AB}_{\mathcal{N}}} + w^k_{\mathtt{AB}}|\end{aligned} \tag{2.24}$$

The probability of success for $\mathcal{M}$ is evaluated as

$$\begin{aligned}
\mathbb{P}[\mathsf{Success}] &= \mathbb{P}[(\mathsf{Success}) \cap (\tau^k_{\mathtt{AB}_{\mathcal{M}}} < 0)] + \mathbb{P}[(\mathsf{Success}) \cap (\tau^k_{\mathtt{AB}_{\mathcal{M}}} \geq 0)] \\
&= \mathbb{P}[(\mathsf{Success})|(\tau^k_{\mathtt{AB}_{\mathcal{M}}} < 0)]\mathbb{P}[\tau^k_{\mathtt{AB}_{\mathcal{M}}} < 0] + \mathbb{P}[(\mathsf{Success}) \cap (\tau^k_{\mathtt{AB}_{\mathcal{M}}} \geq 0)] \\
&\leq \mathbb{P}[\tau^k_{\mathtt{AB}_{\mathcal{M}}} < 0] + \mathbb{P}[(\mathsf{Success}) \cap (\tau^k_{\mathtt{AB}_{\mathcal{M}}} \geq 0)] \\
&< \mathsf{negl}(n) + \mathbb{P}[(\mathsf{Success}) \cap (\tau^k_{\mathtt{AB}_{\mathcal{M}}} \geq 0)] \qquad (2.25)
\end{aligned}$$

In the case where $\tau_{\mathtt{AB}_{\mathcal{M}}} \geq 0$, (2.24) simplifies to

$$\tau^k_{\mathtt{AB}_{\mathcal{M}}} \geq L - |\tilde{\tau}^k_{\mathtt{AB}_{\mathcal{N}}} + w^k_{\mathtt{AB}}|$$

Substituting the least possible value of $\tau^k_{\mathtt{AB}_{\mathcal{M}}}$ into (2.23), it follows that

$$|\tilde{\tau}^k_{\mathtt{AB}_{\mathcal{N}}} + L - |\tilde{\tau}^k_{\mathtt{AB}_{\mathcal{N}}} + w^k_{\mathtt{AB}}| + \tilde{\tau}^l_{\mathtt{BA}_{\mathcal{N}}} + \tau^l_{\mathtt{BA}_{\mathcal{M}}} + \tilde{\tau}_{\mathtt{BB}} + w^l_{\mathtt{BA}}| \leq L - \delta$$

Notice that from the assumptions made in (2.14), (2.20), (2.21), and (2.22), all terms except L and $\tau^l_{\mathtt{BA}_{\mathcal{M}}}$ on the left-hand side of the inequality are negligible compared to L; thus,

$$|L + \tau^l_{\mathtt{BA}_{\mathcal{M}}}| \leq L - \delta$$

Since both L and $L-\delta$ are defined to be positive, the above inequality simplifies to

$$\tau^l_{\mathtt{BA}_{\mathcal{M}}} \leq -\delta$$

where $\delta > 0$. Thus, for $\mathcal{M}$ to succeed in the case where $\tau^k_{\mathtt{AB}_{\mathcal{M}}} \geq 0$, we must have that $\tau^l_{\mathtt{BA}_{\mathcal{M}}} < 0$. As a result

$$\mathbb{P}[(\mathsf{Success}) \cap (\tau^k_{\mathtt{AB}_{\mathcal{M}}} \geq 0)] < \mathsf{negl}(n)$$

Thus, from (2.25)

$$\mathbb{P}[\mathsf{Success}] < \mathsf{negl}(n)$$

Qualitatively, the proof presented here argues that for the adversary to succeed, it needs to either advance the *sync* signal ($\tau_{\mathtt{AB}_{\mathcal{M}}} < 0$), or advance the *response* signal ($\tau_{\mathtt{BA}_{\mathcal{M}}} < 0$). With the use of a secure MAC (or digital signature) and the hardness of superluminal propagation, the adversary can only succeed with a negligible probability.

2.7 Secure Constructions

This section specializes the necessary and sufficient conditions for secure clock synchronization to IEEE 1588 PTP. In addition, it presents an alternative to PTP for wireless synchronization—a compliant synchronization system with GNSS-like signals.

2.7.1 Secure IEEE 1588 PTP

The necessary and sufficient conditions for secure synchronization, as adapted to IEEE 1588 PTP, are as follows:

[T1] Stations `A` and `B` must use an authenticated encryption scheme to prevent $\mathcal{M}$ from generating valid message/tag pairs.

[T2] The difference between the path delays between `A` and `B` and the shortest possible path delays must be negligible as compared to L. For wireless PTP [41,91], this implies communicating over the LOS channel as $L \to 0$.

For traditional wireline PTP, A and B must attempt to communicate over the (nearly) shortest possible path.

[T3] The path delay, which is usually estimated from the RTT measurements, must be accurately known *a priori* for secure synchronization. The RTT measurements must be verified against the expected RTT. This implies that the layover time $\bar{\tau}_{\mathtt{BB}}$ must also be known to A.

Note that in the usual PTP formulation, the path delay is measured and used by the time seeker B. To this end, in the usual formulation A sends the transmit time of the *sync* message and the receipt time of the *delay_req* message (in PTP parlance). Similar conventions may be accommodated in the system model presented in this chapter, wherein A sends the values of $t_{\mathtt{A}}^{\mathtt{A}_k}$, $z_{\mathtt{A}}^{l}$, and $\bar{\tau}_{\mathrm{RTT}}^{kl}$ to B, and the following calculations may be performed and used at B. However, this would only be a cosmetic change and does not affect the arguments in this chapter.

The first security condition has already been proposed in the IEEE 1588-2008 standard. The second condition, however, has not been considered in any of the earlier works in the literature. Following the depiction of *sync* and *response* signal exchange in Fig. 2.2 and the attack strategy outlined in Section 2.5.1, Fig. 2.3 illustrates an example attack against a PTP implementation that does not satisfy the second necessary condition. Notice that the existence of a shorter time path enables $\mathcal{M}$ to advance the *sync* signal relative to the authentic message from A. Subsequently, $\mathcal{M}$ is able to undo the effect of

the advance on the RTT by delaying the *response* signal from B to A. Station A does not measure any abnormality in the RTT, and thus cannot raise an alarm. Meanwhile, synchronization has been compromised at B.

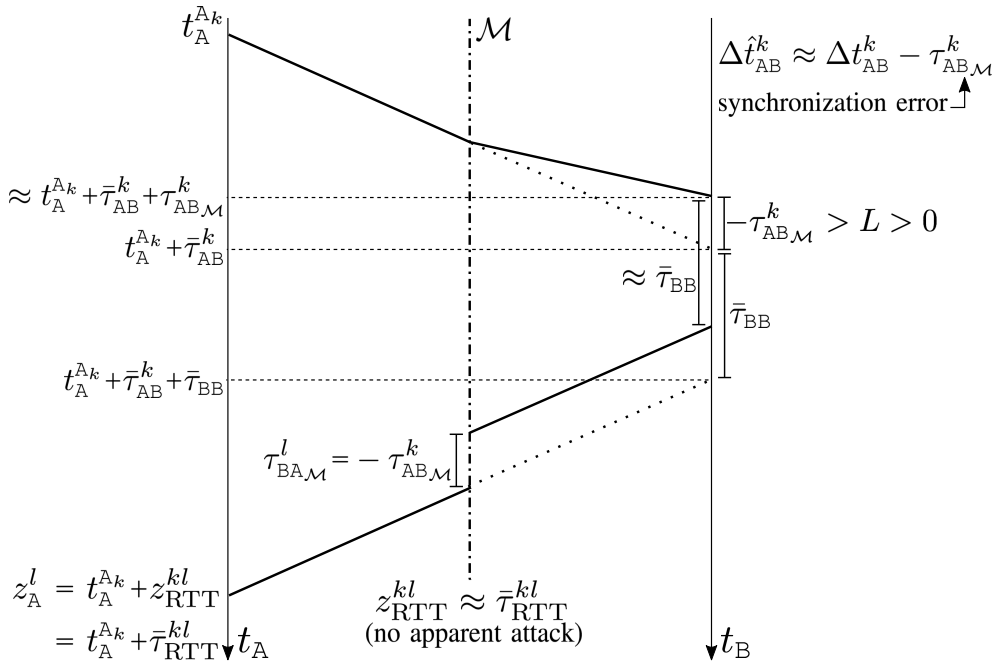


Figure 2.3: Illustration of an example attack against a PTP implementation that violates the second necessary condition.

The third condition is similar to the proposal in [158] of measuring the path delays during initialization and monitoring the delays during normal operation. However, [158] requires that B respond to A with zero delay during initialization to enable measurement of the reference delays. This condition is sufficient, but not necessary for secure synchronization. The system is in fact secure even if B is allowed a fixed layover time. Fig. 2.4 illustrates an example attack against a PTP implementation in violation of the third neces-

sary condition. Note that the uncertainty of the *a priori* estimate of the RTT ($\bar{\sigma}_{\mathrm{RTT}}$) is larger than the alert limit, violating the third necessary condition which requires that the expected RTT be known to within the alert limit (and with much higher accuracy for provable sufficiency). Even though the measured RTT in this case is inconsistent with the expected RTT, it cannot be definitively flagged as an attack since benign variations in the RTT may also have led to the observed RTT.

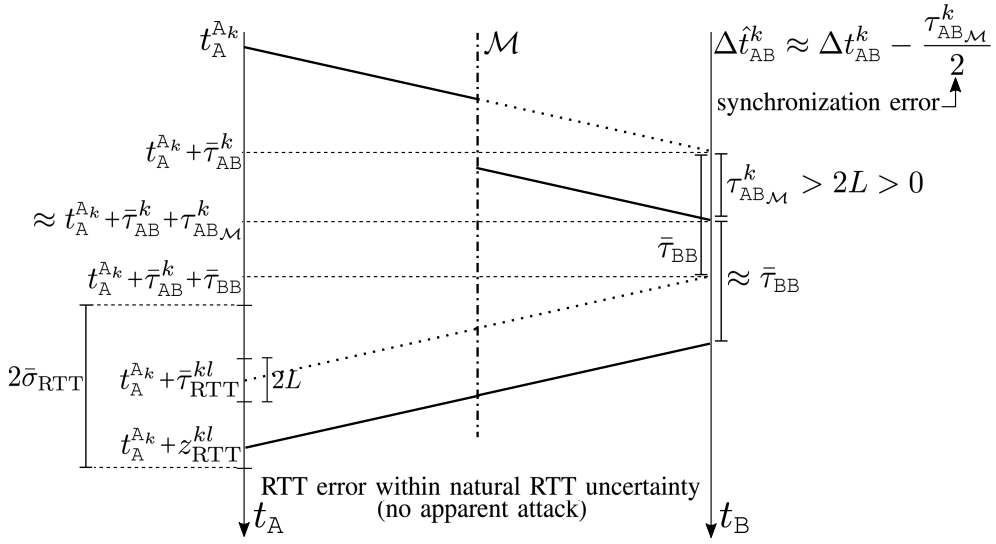


Figure 2.4: Illustration of an example attack against a PTP implementation that violates the third necessary condition.

Interestingly, at first sight, the third security condition in this chapter does not resemble the proposed defense in [9] that enforces an upper bound on the synchronization error accumulated between *sync* messages and recommends that `B` send its timestamps to `A` periodically for verification. As explained next, this condition is in fact equivalent to the condition of known

and measurable RTT, when adapted according to the system model considered in [9].

Note that the requirement of a zero delay in [158], or a short layover time in this chapter, enables A to measure the RTT since the transmit time of the lth feature in $s_{\mathtt{B}}$, that is $t_{\mathtt{B}}^{\mathtt{B}_l}$, can be approximately traced back to A's clock to within the alert limit as $t_{\mathtt{A}}^{\mathtt{A}_k} + \bar{\tau}_{\mathtt{AB}}^{k} + \bar{\tau}_{\mathtt{BB}}$. Enforcing the synchronization error to within L and transmitting B's timestamp to A achieves the same objective for the defense in [9], since the transmit time from B can be traced back to A's clock with the assumed approximate synchronization. Therefore, the proposed countermeasures in [158] and [9] are two different incarnations of the third security condition proposed in this chapter. Of course, the failure of both [158] and [9] to address the second necessary condition makes their proposed defenses vulnerable to an adversary that can communicate along a shorter time path between A and B.

2.7.2 Alternative Compliant System

This section describes an alternative wireless clock synchronization protocol that satisfies the set of necessary and sufficient conditions presented in Section 2.5. The proposed protocol involves bi-directional exchange of GNSS-like pseudo-random codes for continuous clock synchronization, in contrast to discrete packet-based synchronization techniques such as NTP and PTP. It is offered here to illustrate the general applicability of the proposed necessary and sufficient conditions to a range of underlying protocols. Such a protocol

can potentially be applied in two-way satellite time transfer and terrestrial wireless clock synchronization systems for continuous clock synchronization, in contrast to the packet-based discrete synchronization in NTP/PTP.

The time master A and the time seeker B communicate wirelessly over the LOS channel between the nodes. To simplify the analysis, it is assumed that A and B securely share long sequences of pseudo-random bits prior to synchronization. These sequences of bits will later enable generation of unpredictable signals. The pseudo-random sequence for A has the form

$$\boldsymbol{b}_{\mathtt{A}} = \left\{b_{\mathtt{A}}^k\right\}_{k=0}^{N}, \quad b_{\mathtt{A}}^k \in \{0,1\}$$

The pseudo-random code $C_{\mathtt{A}}(t_{\mathtt{A}})$ for A is then generated as

$$C_{\mathtt{A}}(t_{\mathtt{A}}) = 2b_{\mathtt{A}}^k - 1 \text{ for } t_{\mathtt{A}} \in [t_{\mathtt{A}}^{\mathtt{A}_k}, t_{\mathtt{A}}^{\mathtt{A}_{k+1}}), k \in \{0,1,2,\dots\}$$

where $t_{\mathtt{A}}^{\mathtt{A}_k}$ denotes the time according to A at which the start of the kth bit in A's signal is transmitted. The pseudo-random nature of $\boldsymbol{b}_{\mathtt{A}}$ ensures that $C_{\mathtt{A}}(t_{\mathtt{A}})$ has good cross-correlation properties, which enables an accurate measurement of the time-of-arrival of A's signal at B, that is, $\sigma_\epsilon \ll L$. Station A modulates a carrier with the code $C_{\mathtt{A}}$ and transmits a signal $s_{\mathtt{A}}(t_{\mathtt{A}})$ whose complex baseband representation is given as

$$s_{\mathtt{A}}(t_{\mathtt{A}}) = C_{\mathtt{A}}(t_{\mathtt{A}}) \exp\left(j\theta_{\mathtt{A}}(t_{\mathtt{A}})\right)$$

This signal is received at B as

$$\begin{aligned} r_{\mathtt{B}}(t_{\mathtt{A}}, \tau_{\mathtt{AB}}) &= s_{\mathtt{A}}(t_{\mathtt{A}} - \tau_{\mathtt{AB}}) + w_{\mathtt{AB}}(t_{\mathtt{A}}) \\ &= C_{\mathtt{A}}(t_{\mathtt{A}} - \tau_{\mathtt{AB}}) \exp\left(j\theta_{\mathtt{A}}(t_{\mathtt{A}} - \tau_{\mathtt{AB}})\right) + w_{\mathtt{AB}}(t_{\mathtt{A}}) \end{aligned}$$

where all symbols have their usual meanings as detailed in Section 2.4. Station B captures a window $R_{\mathtt{B}}^k$ of $r_{\mathtt{B}}$ and correlates it with a local replica of $C_{\mathtt{A}}$. The result of the correlation enables B to detect the start of the kth bit of $C_{\mathtt{A}}$ in the window, and provides a measurement

$$z_{\mathtt{B}}^k = t_{\mathtt{B}}^{\mathtt{A}_k} + w_{\mathtt{AB}}^k$$

of the time-of-arrival of the kth bit at B. Furthermore, the relationship between the start of the kth bit and $t_{\mathtt{A}}^{\mathtt{A}_k}$ enables B to infer the latter.

If a prior estimate $\bar{\tau}_{\mathtt{AB}}^k$ of $\tau_{\mathtt{AB}}^k$ is available, then B estimates the clock offset $\Delta t_{\mathtt{AB}}^k$ as in (2.4).

Similar to the pseudo-random sequence and code construction for A, B generates its unpredictable code $C_{\mathtt{B}}(t_{\mathtt{B}})$. A and B agree on a one-to-one mapping between $C_{\mathtt{A}}$ and $C_{\mathtt{B}}$ such that B responds with the lth bit of $C_{\mathtt{B}}$ on reception of the start of the kth bit of $C_{\mathtt{A}}$. Furthermore, A and B agree that the start of the lth bit of $C_{\mathtt{B}}$ will have a code-phase offset of $\bar{\tau}_{\mathtt{BB}}$ with respect to the start of the kth bit of $C_{\mathtt{A}}$. Station B transmits the *response* signal as

$$s_{\mathtt{B}}(t_{\mathtt{B}}) = C_{\mathtt{B}}(t_{\mathtt{B}}) \exp\left(j\theta_{\mathtt{B}}(t_{\mathtt{B}})\right)$$

such that

$$t_{\mathtt{B}}^{\mathtt{B}_l} = z_{\mathtt{B}}^k + \bar{\tau}_{\mathtt{BB}}$$

according to the time at B. In true time, the epoch $t_{\mathtt{B}}^{\mathtt{B}_l}$ corresponds to

$$t_{\mathtt{B}}^{\mathtt{B}_l} \rightleftharpoons t_{\mathtt{A}}^{\mathtt{A}_k} + \tau_{\mathtt{AB}}^k + w_{\mathtt{AB}}^k + \tau_{\mathtt{BB}}$$

Station A receives the *response* as

$$r_{\mathtt{A}} = s_{\mathtt{B}}(t_{\mathtt{B}} - \tau_{\mathtt{BA}}) + w_{\mathtt{BA}}(t_{\mathtt{A}})$$

and captures a window of the signal $R^l_{\mathtt{A}}$. A correlates $R^l_{\mathtt{A}}$ with a local replica of $C_{\mathtt{B}}$ to detect the start of the lth bit of $C_{\mathtt{B}}$. This enables A to measure the time-of-arrival

$$z^l_{\mathtt{A}} = t^{\mathtt{B}_l}_{\mathtt{A}} + w^l_{\mathtt{BA}}$$

Moreover, the detection of the lth bit indicates that it was transmitted in response to the receipt of the start of the kth bit of $C_{\mathtt{A}}$. Since A knows the start time of the kth bit as $t^{\mathtt{A}_k}_{\mathtt{A}}$, it measures the RTT as described in (2.6).

Note that the exchange of one-time pad sequences enables the proposed system to satisfy the first security condition. Wireless LOS communication satisfies the second security condition, and the knowledge of the code-phase layover offset enables A to make an accurate prior estimate of the RTT within the alert limit, thereby satisfying the third security condition. Thus, the proposed system complies with all three necessary and sufficient conditions for secure clock synchronization.

2.8 System Simulation

This section presents a simulation study of a secure clock synchronization model operating over a simplistic channel model. Unlike the abstract treatment of delays in the security derivations presented earlier, the simulation is carried out with models of delays experienced by the synchronization

messages over a real channel. This study also expounds the interplay between slave clock stability, security requirements, attack models, and attack detection thresholds that must be determined in a practical synchronization system. The channel and attack models developed in this simulation are not comprehensive. Rather, relatively simple models are considered to clearly demonstrate the underlying principles. More sophisticated channel and attack models can similarly be analyzed by following the outline of this simulation.

2.8.1 Channel Model

The simulated system resembles a traditional local area network, and is schematically depicted in Fig. 2.5. As before, `A` and `B` are the time master and seeker stations, respectively. The messages between these stations pass through a series of N routers. Each router is under network traffic loading generated by the nodes labeled `T`. The routers perform simple packet forwarding, i.e., no cryptographic operations or complex payload modifications are performed. Each router transmits the queued packets at a service rate of 1 Gbps. Each network packet is assumed to have a size of 1542 bytes. The MITM adversary $\mathcal{M}$ maliciously inserts itself along the communication path between `A` and `B`.

The *sync* and *response* packets from `A` and `B` experience processing and queueing delay at each router, and propagation/link delay between routers. Queueing delay is the duration for which the packet is buffered in the router before it can be transmitted. Processing delay is the time taken by the router

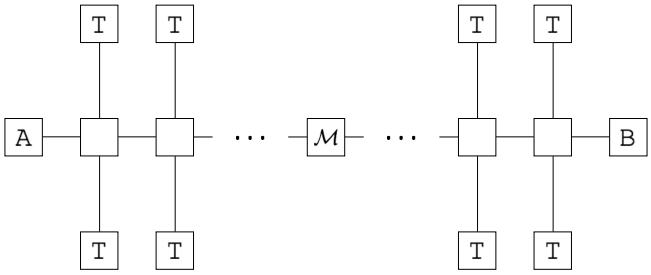


Figure 2.5: Schematic diagram of the network topology considered in this section.

to process the packet header, for example, to determine the packet's destination. Since the routers in this simulation perform simple packet forwarding, the processing delay is negligible as compared to the queueing delay [126]. The propagation/link delay is also insignificant for local networks because the propagation speed is a comparable fraction of the speed of light. Thus, only the queueing delay significantly contributes to the overall channel delay variations.

Let the network idle probability for a particular router, denoted by ρ, be defined as the probability of the router queue being empty at a randomly chosen time instant. Since the synchronization packets are delay-sensitive, the routers in this simulation implement non-preemptive priority scheduling for synchronization packets when the queue is not empty. This means that on arrival of a *sync* or *response* packet, the router is allowed to complete the transmission of the data packet currently being serviced, if any, but is required to service the delay-sensitive packet before the other network data in the queue. Since the time period between consecutive *sync-response* pairs is quite large as compared to the RTT for a given pair, it is assumed that a

router never has more than one delay-sensitive packet in its queue. Under such scheduling, the delay experienced by the timing messages is best modeled as follows: with probability ρ, the total router delay is zero, and with probability $(1-\rho)$ the total router delay is uniformly distributed between zero and the maximum time to service a packet of length 1542 bytes ($1542 \times 8 \times 2^{-30} \approx 11.49$ microseconds for a Gigabit router).

Given the above channel specifications and values for N and ρ, it is possible to perform a Monte Carlo simulation to obtain the anticipated RTT $\bar{\tau}_{\mathrm{RTT}}$, which is taken to be the empirical mean of the RTT measurements in the simulation, and the associated standard deviation $\bar{\sigma}_{\mathrm{RTT}}$. As shown in Fig. 2.6, in case of a single *sync-response* pair measurement, the RTT has an empirical mean of 80.34 microseconds and an empirical standard deviation of 17.09 microseconds with $N = 10$ and $\rho = 0.3$. Observe that even for a relatively small N, the empirical distribution approaches the Gaussian shape, but has slightly heavier tail on the higher end of the delay. The distribution for mean of batches of 10 observations has a smaller empirical standard deviation of 5.41 microseconds.

2.8.2 System and Security Requirements

The clock at the time seeker `B` drifts with respect to the true time clock at `A` unless corrected by a *sync* message from `A`. As before, let L denote the alert limit for the system. Let T denote a time duration over which a perfectly synchronized clock at `B` at the beginning of the duration, absent an adversary,

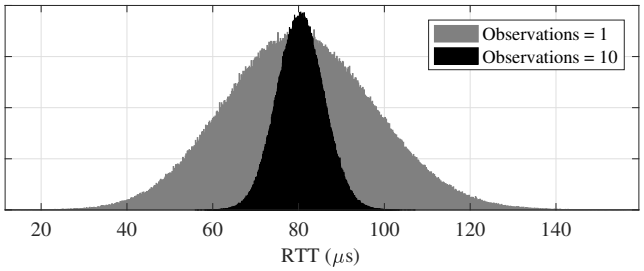


Figure 2.6: Empirical distribution of the RTT of *sync-response* pairs through a network of $N = 10$ routers with network idle probability of $\rho = 0.3$. The light-shaded histogram shows the empirical distribution of the RTT of a single *sync-response* pair. The dark-shaded histogram shows the corresponding distribution for the mean of batches of 10 observations of the RTT.

drifts more than $L_{\mathcal{N}}$ for some $L_{\mathcal{N}} < L$ with a probability smaller than an acceptably small bound $\mathbb{P}_{\epsilon}$.

In the system under simulation, the clock offset for `B` is estimated and corrected for every T seconds. By definition of T, it holds that if the clocks at `A` and `B` are perfectly synchronized after every T seconds, then the natural drift envelope of `B`'s clock does not exceed L with an unacceptably high probability. Define

$$L_{\mathcal{M}} \triangleq L - L_{\mathcal{N}}$$

Observe that if an adversary is able to introduce a synchronization error larger than $L_{\mathcal{M}}$, then the system is compromised since the natural drift of the clock at `B` could potentially lead to a clock offset greater than L before the next synchronization interval, with a probability greater than $\mathbb{P}_{\epsilon}$. Thus, `A` must flag any adversarial delay greater than $L - L_{\mathcal{N}}$ with probability higher than a desired detection probability, denoted by $\mathbb{P}_{\mathrm{D}}$. It is worth noting that this

practical complication of the magnitude of L_{N} was abstracted in the sufficiency proof, where the threshold was set to $L - \delta$ for $\delta > 0$.

In general, A makes multiple measurements of the RTT between A and B over the time period T. As shown in Fig. 2.6, the mean of multiple observations over T has a distribution with a smaller standard deviation as compared to that of a single observation. In the simulated system, if no attack is detected, A updates $\bar{\tau}^k_{\mathtt{AB}}$ every T seconds based on the empirical mean of the RTT measurements made over that period. Note that even though $\bar{\tau}^k_{\mathtt{AB}}$ is updated based on the measurements, no updates are applied to $\bar{\tau}_{\mathrm{RTT}}$ and $\bar{\sigma}_{\mathrm{RTT}}$, which are predetermined by simulation or measurements under a secure calibration campaign.

The empirical mean of the measured RTT is taken as the test statistic to detect an attack. For the attack model detailed next, it can be shown that this test statistic becomes optimal for large values of N [161].

2.8.3 Attack Model

The synchronization system considered in this simulation complies with the necessary security conditions presented in this chapter. Consequently, the adversary $\mathcal{M}$ is unable to advance the *sync* or *response* messages, and can only increase the RTT measured by A relative to the authentic RTT. This simulation considers a simple adversary model that introduces a fixed delay in the measured RTT. In order to conceal its presence while compromising synchronization with appreciable probability, $\mathcal{M}$ introduces a delay of $L_{\mathcal{M}} + \xi$

seconds for some small $\xi > 0$.

Let H_0 denote the null hypothesis (no attack), and H_1 denote the alternative hypothesis. Under H_0, the measured RTT at `A` is drawn from the distribution that was used to calibrate/simulate the channel delay distribution, while under H_1, the measured RTT is drawn from a distribution that is shifted from the calibration distribution by $L_{\mathcal{M}}+\xi$. This is visually depicted in Fig. 2.7. Given a detection threshold λ, the dark-shaded region in Fig. 2.7 denotes the probability of false alarm, $\mathbb{P}_{\mathrm{F}}$, while the light-shaded region denotes the probability of missed detection $(1 - \mathbb{P}_{\mathrm{D}})$. In observing Fig. 2.7, it might be argued, and holds true, that a reasonable attacker may introduce noise in the introduced delay to inflate the width of the distribution under H_1 and thereby decrease the probability of detection of an attack. However, in that case, the empirical mean test statistic is no longer optimal. Instead, `A` would incorporate the observed variance of the RTT in its test statistic in addition to the empirical mean. In short, the attack model in this simulation is not comprehensive, as explained previously. For a more sophisticated treatment of sensor deception and protection techniques, the reader may refer to [21].

2.8.4 Simulation

The system and attack described above have been simulated with $N = 10$ and $\rho = 0.3$ for all routers. The adversarial delay $L_{\mathcal{M}} + \xi$ is set to 10 microseconds, and the required probability of detection $\mathbb{P}_{\mathrm{D}}$ is set to 0.999. The number of RTT observations made in time T are varied between 1 and 200.

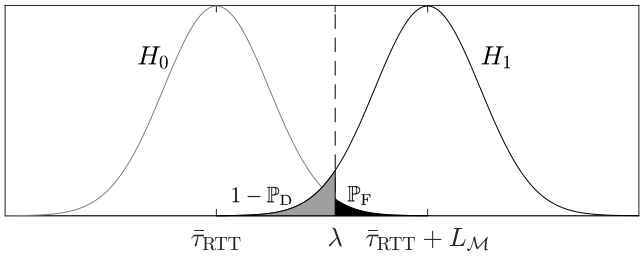


Figure 2.7: Representation of the distributions under H_0 and H_1 along with the detection threshold and the associated $\mathbb{P}_{\mathrm{F}}$ and $\mathbb{P}_{\mathrm{D}}$.

Given the number of observations, and a required $\mathbb{P}_{\mathrm{D}}$, the system is simulated under H_1 for 10^6 detection epochs and the maximum possible detection threshold λ that satisfies the detection probability is obtained. Subsequently, the system is simulated under H_0 and the number of test statistics exceeding the threshold λ are recorded. The frequency of such epochs is reported as the probability of false alarm $\mathbb{P}_{\mathrm{F}}$.

Fig. 2.8 shows the above procedure for 80 RTT measurements made per test statistic. In this case, λ is obtained to be 84.53 microseconds and the corresponding $\mathbb{P}_{\mathrm{F}}$ is 1.59%. Fig. 2.9 shows a log-log plot of $\mathbb{P}_{\mathrm{F}}$ as a function of the number of observations made per test statistic. When the number of observations is greater than 160, no false alarms were observed with 10^6 trials. For the given channel delay variation statistics, the probability of false alarm is very high for small number of observations per decision epoch since the threshold λ that must be set to detect an attack with the required $\mathbb{P}_{\mathrm{D}}$ is large in comparison to the minimal delay that the adversary must introduce to compromise synchronization ($L_{\mathcal{M}}$). For a more stable channel, such as a wireless or PTP-aware channel, fewer measurements per decision epoch would

suffice.

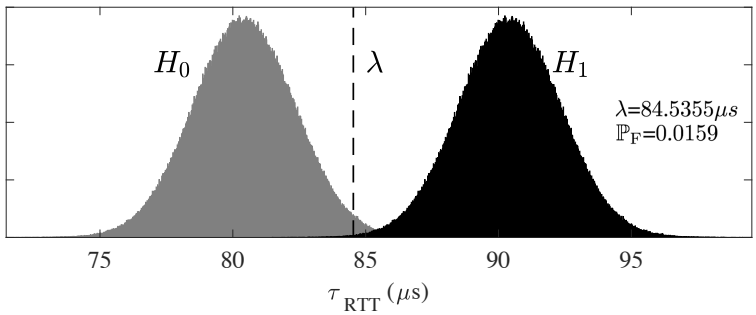


Figure 2.8: Distribution of the test statistic under H_0 and H_1 for 80 RTT measurements per decision epoch. ($N = 10$, $\rho = 0.3$, $L_{\mathcal{M}} + \xi = 10\mu$s, $\mathbb{P}_D = 0.999$)

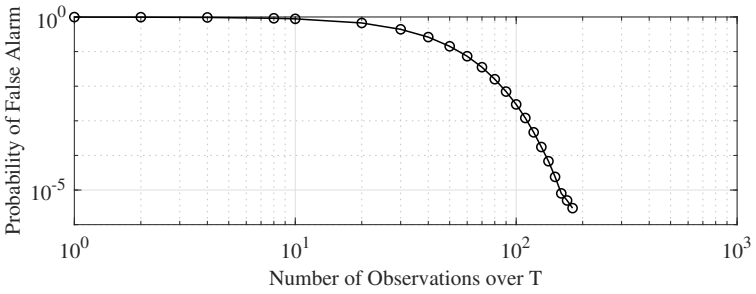


Figure 2.9: Probability of false alarm as a function of number of observations per decision epoch. ($N = 10$, $\rho = 0.3$, $L_{\mathcal{M}} + \xi = 10\mu$s, $\mathbb{P}_D = 0.999$)

2.8.5 Practical Implications

Section 2.6.1 makes fairly remarkable assumptions about the synchronization system to show provably secure time transfer. For instance, it requires that errors in the *a priori* estimate of the RTT of the timing messages be negligible compared to the alert limit. Nonetheless, as shown in this section, for a given channel with bounded delay variations and a given slave clock, some

level of security guarantee can be made for a synchronization system that satisfies the necessary and sufficient conditions presented herein. For concreteness, consider a system that requires an alert limit of $L = 100$ microseconds and a slave clock that drifts no more than $L_{\mathcal{N}} = 50$ microseconds over a period of $T = 1$ second with acceptably high probability $(1 - \mathbb{P}_\epsilon)$. Then, for $N = 10$ and $\rho = 0.3$, if `A` makes 10 RTT measurements over 1 second, the empirical mean test statistic is distributed as the dark-shaded distribution in Fig. 2.6 with a standard deviation of ≈ 5.4 microseconds. For $L_{\mathcal{M}} = L - L_{\mathcal{N}} = 50$ microseconds, a threshold of $\approx \bar{\tau}_{\mathrm{RTT}} + 30$ microseconds will yield a missed detection rate of approximately 1 in 15000, and a false alarm rate of approximately 3.5 in 1 million. With a more stable slave clock or more measurements per second, these probabilities can be made more favorable.

Another important concern that has not been addressed in the simulation is that of the incorporation of cryptographic constructs in the synchronization protocol. The encryption and decryption algorithms are often complex and take non-negligible processing time to execute. However, note that at `A`, the *sync* message is timestamped *after* the encryption process, and thus the time taken for encryption is inconsequential. At `B`, it is important to concede that the decryption of the *sync* message and the encryption of the *response* message cannot be assumed to happen instantaneously. This has been accounted for by allowing the layover time $\bar{\tau}_{\mathtt{BB}}$ for the cryptographic processes to execute. Once again, the receipt timestamp of the *response* message at `A` is applied *before* the decryption process, and hence the decryption time at `A` is

inconsequential. Thus, compliance with the first security condition must not pose significant practical challenges.

2.9 Conclusions

A fundamental theory of secure clock synchronization was developed for a generic system model. The problem of secure clock synchronization was formalized with explicit assumptions, models, and definitions. It was shown that all possible one-way clock synchronization protocols are vulnerable to replay attacks. A set of necessary conditions for secure two-way clock synchronization was proposed and proved. Compliance with these necessary conditions with strict upper bounds was shown to be sufficient for secure clock synchronization, which is a significant result for provable security in critical infrastructure. The general applicability of the set of security conditions was demonstrated by specializing these conditions to design a secure PTP protocol and an alternative secure two-way clock synchronization protocol with GNSS-like signals. Results from a simulation with models of channel delays were presented to expound the interplay between slave clock stability, security requirements, attack models, and attack detection thresholds.

Chapter 3

World-wide GNSS Interference Monitoring from Low-Earth Orbit

3.1 Abstract

Observation of terrestrial GNSS interference (jamming and spoofing) from low-earth orbit (LEO) is a uniquely effective technique for world-wide monitoring of hostile and contested GNSS environments, and for estimating the locations of interference sources. Such details are useful for situational awareness, interference deterrence, and for developing interference-hardened GNSS receivers. This chapter explores the performance of LEO interference monitoring using receiver-reported carrier-to-interference-and-noise ratio (CINR) and presents the results of a three-year study of global interference. Using CINR data from a GNSS receiver aboard the International Space Station (ISS), this study confirms the presence of reported interference activity

This chapter is a subset of: Matthew J. Murrian, Lakshay Narula, Peter A. Iannucci, Scott Budzien, Brady W. O'Hanlon, Steven P. Powell, and Todd E. Humphreys. GNSS interference monitoring from low Earth orbit. *Navigation, Journal of the Institute of Navigation*, 2020. Submitted for review. The material presented in this chapter only includes the contributions made by the author of this dissertation.

in Syria and Libya, and uncovers the existence of hitherto unknown ongoing GNSS interference in mainland China.

3.2 Introduction

Terrestrial GNSS interference activity has grown more widespread and sophisticated over recent years. Conspicuous GNSS jamming or spoofing has occurred, or is ongoing, at urban and coastal sites around the globe [3, 27, 30, 137]. Given the dependence of critical infrastructure and safety-of-life systems on GNSS [71, 122, 139, 165], there is great interest in detecting, characterizing, and localizing sources of interference.

Space-based observation of terrestrial GNSS interference is attractive for several reasons. Most obviously, it offers world-wide coverage: moderately-powerful terrestrial interference sources anywhere on the globe can be detected by LEO satellites multiple times per day, making it possible to maintain a common operating picture of world-wide GNSS interference. A single LEO-based sensor is sufficient to characterize the strength, spectral properties, structural content, and even the location of terrestrial interference sources, provided a Doppler time history can be extracted from a carrier component of the interference signal [106]. For signals from which no carrier can be isolated, multiple synchronized LEO-based sensors can employ time- and frequency-difference-of-arrival (TDOA and FDOA) techniques to infer the source's location [20, 21].

Another simple and effective interference detection test can be formulated solely from the standard carrier-to-noise ratio observable, C/N_0, pro-

duced by a GNSS receiver. The presence of an interference source in the GNSS radio-frequency band reduces the receiver-reported C/N_0, and in the worst case, leads to a complete loss of signal tracking. This chapter presents the results of a three-year study of terrestrial GNSS interference as observed through the reported C/N_0 from a software-defined GNSS receiver operating since February 2017 on the ISS. The FOTON receiver, developed by The University of Texas at Austin (UT) and Cornell University, is part of a larger science experiment called GPS Radio Occultation and Ultraviolet PhotometryColocated (GROUP-C), an unclassified experiment aboard the ISS that is part of the Space Test ProgramHouston Payload 5 (STP-H5) payload.

The FOTON receiver is a science-grade spaceborne dual-frequency (GPS L1 and L2) GNSS receiver [87]. Three levels of FOTON data are available for interference analysis: (1) raw 5.7 Msps intermediate frequency (IF) samples output by the FOTON front-end's analog-to-digital converter, (2) 100-Hz data-modulation-wiped complex correlation products, and (3) 1-Hz standard GNSS observables. The analysis presented in this chapter only makes use of the 1-Hz observables, which, in addition to the C/N_0, includes the receiver's position and velocity, and flags indicating the health of the signals being tracked.

The two contributions of this chapter are (1) it provides an analysis of expected performance for terrestrial GNSS interference monitoring from LEO with C/N_0 observables, and (2) it presents the results of a three-year study of global GNSS interference monitoring from the ISS. This chapter is a subset of the work presented in [106]. In particular, only direct contributions of the

author of this dissertation have been included in this chapter.

No prior public literature explores the use of a space-borne GNSS receiver for monitoring terrestrial GNSS interference, despite increasing concern over such interference [65, 122, 123, 164].

3.3 Interference Detection Performance via C/N_0 Monitoring

This section explores the potential performance of LEO GNSS interference monitoring based on receiver-reported C/N_0 in terms of detection sensitivity.

Detection of GNSS interference can be broadly classified as operating at the pre- or post-correlation stage within a GNSS receiver [28]. Pre-correlation detection is much less sensitive, but works for all signals with power falling in the band of interest. Post-correlation detection can only be applied to structured interference with a known waveform, but due to processing gain, is much more sensitive. What follows is a sensitivity analysis of the C/N_0-based interference detection technique, which operates at the pre-correlation stage (with respect to the interference signal).

A simple and effective pre-correlation interference detection test can be formulated solely from the standard carrier-to-noise ratio observable, C/N_0, produced by a GNSS receiver. In the presence of interference, C/N_0 actually measures the carrier-to-interference-and-noise ratio, CINR. Let C be the received authentic signal power for a particular satellite-and-signal combination

[e.g., the GPS L1 C/A signal corresponding to pseudo-random number (PRN) code 4], N_0 be the (approximately flat) receiver thermal noise power density near the frequency band of interest, and I_0 be the spectrally-flat-equivalent interference noise power density, whose relationship with the actual interference power spectrum is described in [65]. Then CINR is defined as

$$\text{CINR} \equiv \frac{C}{N_0 + I_0} \tag{3.1}$$

When compensated for satellite- and receiver-side antenna gain patterns and for spreading loss along the satellite-to-receiver path, and absent signal blockage, strong scintillation, and "flex-power" satellite power adjustments, CINR variations are primarily driven by multipath, which is characterized by a log-normal distribution [164]. Let $\boldsymbol{z}$ be a vector of CINR measurements expressed in dB for a particular frequency band, with predictable variations due to antenna gain pattern and spreading loss removed. A hypothesis test for interference can be formulated in terms of the common decrease in the elements of $\boldsymbol{z}$ due to an increase in I_0. In particular, the distribution of $\boldsymbol{z}$ under the null (H_0) and alternate (H_1) hypotheses may be modeled as

$$H_0 : \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{\mu}, P) \tag{3.2a}$$

$$H_1 : \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{\mu} - \delta\mathbf{1}, P) \tag{3.2b}$$

where $\boldsymbol{\mu} \in \mathbb{R}^{n_z}$, $P \in \mathbb{R}^{n_z \times n_z}$, $\mathbf{1}$ denotes an all-ones column vector of the same length as $\boldsymbol{\mu}$, and $\delta > 0$ is the amount in dB by which all CINR values drop due to interference under H_1.

The model in (3.2) conservatively assumes that $\boldsymbol{z}$'s covariance matrix, P, is identical for H_0 and H_1. In practice, although the receiver's multipath environment remains unchanged from H_0 to H_1, interference sources can cause time variations in I_0 that inflate P in the positive definite sense. But because the magnitude of increase in P is impossible to know *a priori*, the less-sensitive model presented above is assumed.

The model in (3.2) is a special case of the general Gaussian problem for which the likelihood ratio test can be reduced to [155]

$$l(\boldsymbol{z}) = \mathbf{1}^T P^{-1} \boldsymbol{z} \underset{H_1}{\overset{H_0}{\gtrless}} \nu \tag{3.3}$$

where $l(\boldsymbol{z})$ is the test's (sufficient) detection statistic. This test is optimal despite δ being unknown *a priori* because $l(\boldsymbol{z})$ is independent of δ (i.e., the test is uniformly most powerful with respect to δ). Note that P may not be diagonal because the elements of $\boldsymbol{z}$ may be correlated through dependence on the spacecraft attitude or because $\boldsymbol{z}$ may contain multiple elements for the same satellite-signal pair taken over a sliding window of time.

As a linear transformation of a Gaussian vector, $l(\boldsymbol{z})$ is itself Gaussian. Hence, the performance of the test in (3.3) can be completely characterized by the normalized distance between the means of $l(\boldsymbol{z})$ under H_0 and H_1:

$$d \equiv \frac{\mathbb{E}[l|H_0] - \mathbb{E}[l|H_1]}{\sqrt{\mathrm{Var}(l|H_0)}} = \delta\sqrt{\mathbf{1}^T P^{-1}\mathbf{1}} \tag{3.4}$$

Fig. 3.1 shows how the performance improves with increasing d.

If the CINR measurements in $\boldsymbol{z}$ are taken at a single epoch of time, and if the effect of multipath on each measurement is only weakly coupled through

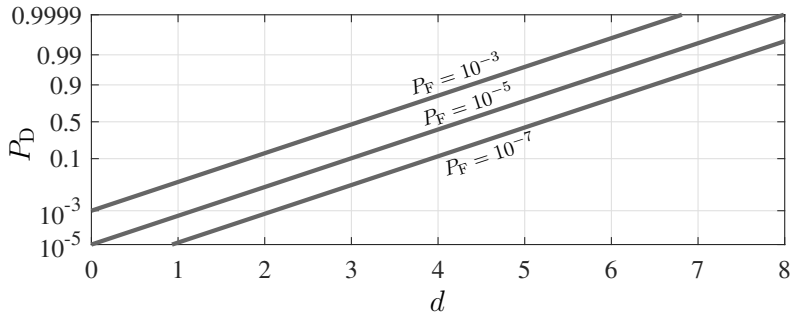


Figure 3.1: Detection probability for the test in (3.3) as a function of d for three different values of false alarm probability.

the spacecraft attitude, then P may be modeled as diagonal. In the simplest case, $P = \sigma_z^2 I$ and d reduces to

$$d = \delta\sqrt{n_z}/\sigma_z \tag{3.5}$$

For the FOTON receiver on the ISS, the ISS's extended shape and large solar panels create an unfavorable multipath environment, resulting in a relatively high $\sigma_z \approx 1.5$ dB. More compact LEO satellites such as the main sounding rocket payload in [87] enjoy $\sigma_z < 1$ dB.

Approximate LEO interference detection sensitivity in the L1 GNSS band using only CINR measurements can be calculated by assuming $\sigma_z = 1$ dB and $n_z = 15$, which are reasonable parameters for a single-epoch test, a horizontally-oriented hemispherical-gain antenna, and full constellations of GPS, Galileo, and BDS III satellites. From (3.5) and Fig. 3.1, a drop in CINR of $\delta > 1.4$ dB is required at $P_F = 10^{-5}$ to yield $P_D > 0.9$. Conservatively assuming that the interference power is spread evenly across the 4-MHz bandwidth covering the most-widely-used civil L1 GNSS signals, then

$I_0 = P_\mathrm{I} - 66$ dBW/Hz, where P_I is the received interference power in dBW. Assuming $N_0 = -204$ dBW/Hz, a CINR drop by $\delta = 1.4$ dB implies $P_\mathrm{I} = -142$ dBW. Denote spreading loss by L dB, receiver antenna gain by G_r dB, and interference source effective isotropic radiated power (EIRP) by P_EIRP dBW. Then

$$P_\mathrm{EIRP} = P_\mathrm{I} - G_r + L \tag{3.6}$$

Spreading loss at L1 from the surface along the shortest distance to a typical LEO altitude of 400 km is $L = 148.5$ dB. Then, supposing $G_r = 3$ dB, the minimum EIRP of an interference source detectable solely from CINR measurements with $P_\mathrm{F} \leq 10^{-5}$ and $P_\mathrm{D} > 0.9$ is approximately $P_\mathrm{EIRP} = 3.5$ dBW.

3.4 LEO Interference Survey via C/N_0 Observables

The 1-Hz standard GNSS observables have been logged from the FOTON receiver nearly continuously since early 2017. These data facilitate a world-wide survey of strong GNSS interference.

The carrier power C of an authentic signal can be modeled as a function $C(j, f, r_{sr}, z_s, z_r)$, where j is the GNSS satellite identifier (SV ID), f is the frequency band (L1 or L2), r_{sr} is the range between the GNSS satellite antenna and the ISS FOTON antenna, z_s is the angle between the satellite boresight direction and the direction to the ISS antenna (i.e., the satellite antenna zenith angle), and z_r is the angle between the ISS antenna boresight direction and the direction to the satellite (the receiver antenna zenith angle). As discussed in Section 3.3, a hypothesis test based on the receiver-reported CINR can

be designed to detect whether (H_1) or not (H_0) the receiver is experiencing interference. Setting up this test based on the FOTON data presents two additional challenges:

[T1] The hypothesis test requires that the statistics $\mathbb{E}[l|H_0]$ and $\mathrm{Var}(l|H_0)$ be known. This requires a "calibration" process where the receiver is known to be in an interference-free environment. Since this knowledge cannot be guaranteed in a practical setting, an assumption needs to be made. This section assumes the receiver reports interference-free data (consistent with H_0) when the ISS is over deep ocean bodies.

[T2] Modeling of interference-free C/N_0 is further complicated by the ISS's local multipath environment. The ISS antenna is flanked by solar panels that move with respect to the FOTON antenna, causing a non-stationary signal obstruction and multipath environment. Nevertheless, a zenith angle window $z_r \in [0°, 15°]$ is known to be free of obstructions. Only the signals received in this window are considered for interference detection in this chapter's analysis.

To maximize the sensitivity of the hypothesis test, one must eliminate or model all non-interference sources of variation in the reported CINR. First, observe that the dependence of C on r_{sr} can be easily removed by compensating for the free space path loss:

$$\hat{C}(j, f, z_s, z_r) = C(j, f, r_{sr}, z_s, z_r) \times \left(\frac{4\pi r_{sr} f}{c}\right)^2$$

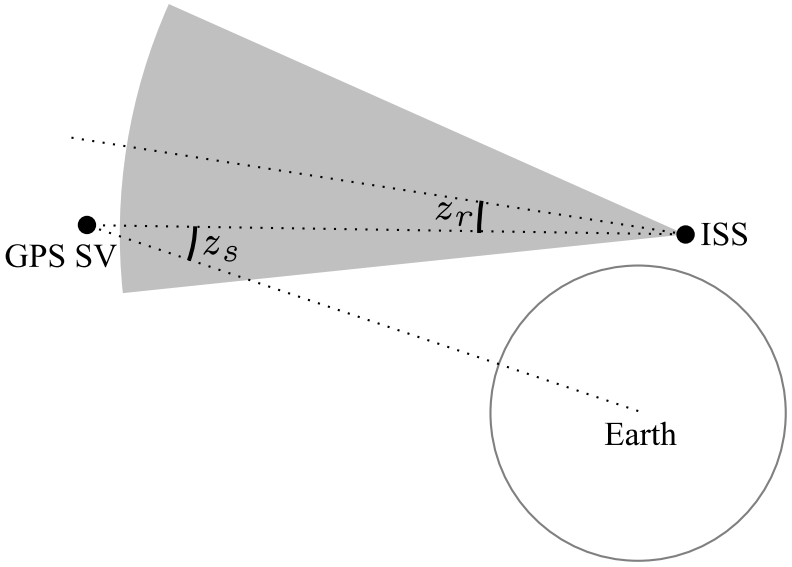


Figure 3.2: For receiver zenith angle $z_r \leq 15^\circ$ (within the gray region), the satellite zenith angle z_s is restricted between $14.2^\circ \leq z_s \leq 15.2^\circ$

Effectively, this eliminates the variation in CINR due to different GNSS satellites being at different distances from the receiver.

Furthermore, it turns out that given the restriction of $z_r \in [0^\circ, 15^\circ]$, the geometry between GNSS satellites and the ISS is restricted such that $z_s \in [14.2^\circ, 15.2^\circ]$ (see Fig. 3.2). This implies that the dependence of $\hat{C}$ on z_s may be ignored, since the GNSS antenna gain pattern can be assumed to be constant over $\pm 0.5^\circ$. In other words, $\hat{C}$ is only a function of the SV ID j, the frequency band f, and the receiver zenith angle z_r.

To set up the hypothesis test, the mean and variance of ISS-reported range-compensated-CINR values $\hat{C}/N_0$ collected over deep ocean regions are maintained as control data (H_0) in a three-dimensional grid of SV ID, fre-

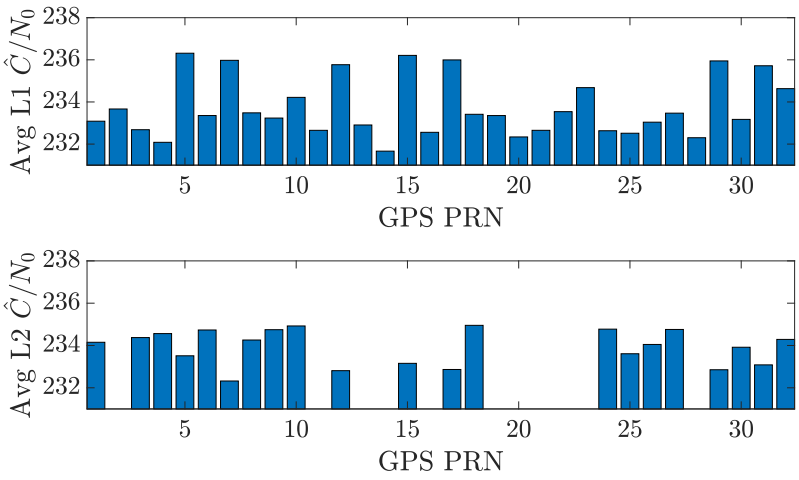


Figure 3.3: Average interference-free range-compensated carrier-to-noise ratio $\hat{C}/N_0$ at L1 (top panel) and L2 (bottom panel) frequencies for GPS satellites over 3 years of collected data. Notice that some SVs are up to 4 dB more powerful than others at GPS L1. The blank bars in the bottom panel correspond to GPS satellites without an L2 signal. Compensating for this effect increases the sensitivity of the hypothesis test.

quency band, and receiver zenith angle.

Fig. 3.3 shows the average interference-free range-compensated carrier-to-noise ratio $\hat{C}/N_0$ at L1 and L2 frequencies for GPS satellites over 3 years of collected data. Notice that some GPS L1 SVs are up to 4 dB more powerful than others. Compensating for this effect increases the sensitivity of the hypothesis test by reducing the uncertainty of the test statistic under H_0.

The test sensitivity is further improved by compensating for the gain pattern of the GPS antenna at the ISS. Fig. 3.4 shows the average interference-free mean-adjusted $\hat{C}/N_0$ at L1 and L2 frequencies as a function of the ISS zenith angle z_s. Mean adjustment removes the effect of SV-specific power

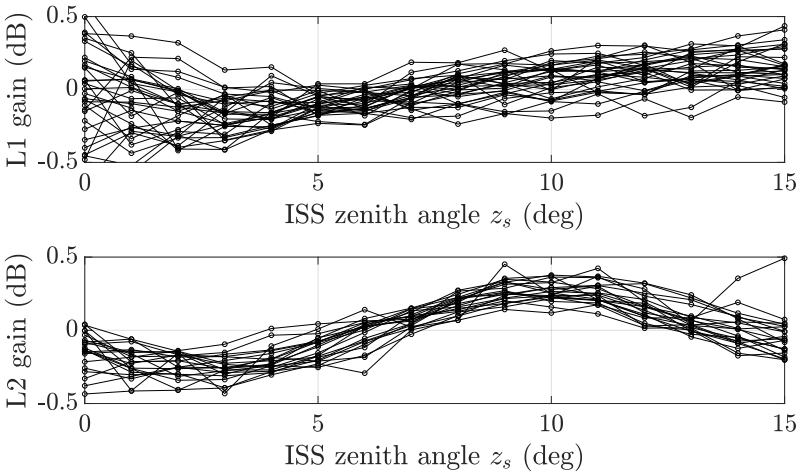


Figure 3.4: Average interference-free mean-adjusted $\hat{C}/N_0$ at L1 (top panel) and L2 (bottom panel) frequencies as a function of the ISS zenith angle z_s. Each line in the chart corresponds to a GPS SV, and provides an estimate of the gain pattern of the GPS antenna at the ISS. The pattern is clear for GPS L2, with 1 dB peak-to-peak gain variation. The L1 pattern is flatter, but shows slightly larger gain for larger zenith angles. Compensating for this effect increases the sensitivity of the hypothesis test.

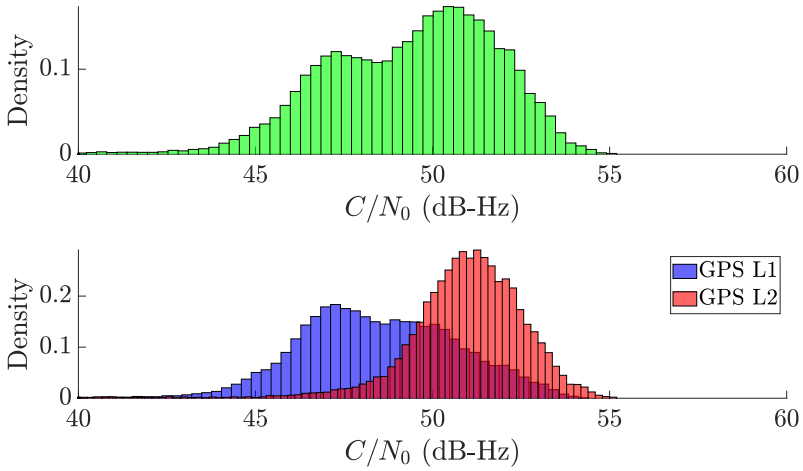


Figure 3.5: Top panel shows a histogram of the receiver-reported interference-free carrier-to-noise ratio C/N_0 for both L1 and L2. Bottom panel shows separate histograms of C/N_0 at the two frequencies. Notice that the histograms are narrower when separated for the two frequencies. Narrower histograms increase the sensitivity of the hypothesis test, i.e., weaker interference is detectable if the interference-free C/N_0 is predictable.

variation observed in Fig. 3.3. The trend in Fig. 3.4 provides an estimate of the gain pattern of the ISS GPS antenna. The pattern is clear at GPS L2, with 1 dB peak-to-peak gain variation. The L1 gain pattern is less pronounced, but shows slightly larger gain for larger zenith angles.

To demonstrate the effect of compensating for the various factors described above, Figs. 3.5, 3.6, and 3.7 show the histograms of the carrier-to-noise ratio data from a single day after the various corrections are applied. The top panel in Fig. 3.5 shows a histogram of the receiver-reported interference-free carrier-to-noise ratio C/N_0 for both L1 and L2 before any processing. The bottom panel shows separate histograms of C/N_0 at the two frequencies. Notice that the histograms are narrower when separated for the two frequencies,

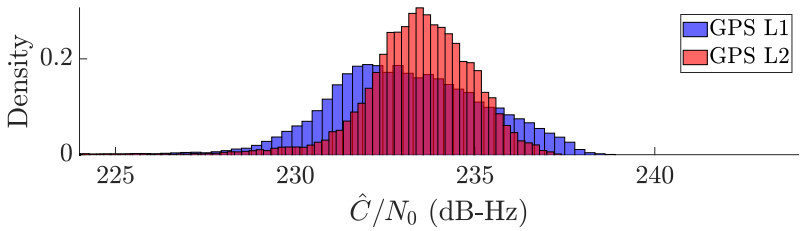


Figure 3.6: Histogram of the interference-free range-compensated carrier-to-noise ratio $\hat{C}/N_0$ for GPS L1 (in blue) and GPS L2 (in red). Notice that when compared to Fig. 3.5, compensating for range reduces the uncertainty under H_0, thus increasing the sensitivity of the hypothesis test.

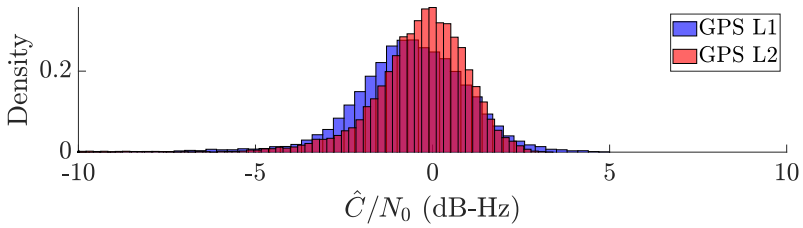


Figure 3.7: Histogram of the interference-free carrier-to-noise ratio after compensating for range to the satellite, GPS SV ID, frequency band, and ISS zenith angle for GPS L1 (in blue) and GPS L2 (in red). Notice that when compared to Figs. 3.5 and 3.6, compensating for the above factors reduces the uncertainty under H_0, thus increasing the sensitivity of the hypothesis test.

increasing the sensitivity of the hypothesis test.

After compensating for the path loss between the GPS SV and the ISS, the resulting histograms for GPS L1 and L2 are shown in Fig. 3.6. Notice that when compared to Fig. 3.5, compensating for range reduces the uncertainty under H_0. Finally, Fig. 3.7 shows the histogram of the interference-free carrier-to-noise ratio after further compensation for the GPS SV ID, the frequency band, and the ISS zenith angle. When compared to Figs. 3.5 and 3.6, compensating for the above factors further reduces the uncertainty under H_0.

The hypothesis test is performed on the test statistic derived from $\hat{C}/N_0$ values that fall within $z_r \in [0°, 15°]$. The test is performed separately for the L1 and L2 bands since the interference characteristics are frequency dependent. If the reported test statistics falls below $\mathbb{E}[l|H_0] - 3\sqrt{\mathrm{Var}(l|H_0)}$, the receiver is declared to be under interference. This threshold respects a P_{F} of approximately 1.35×10^{-3}. Fig. 3.8 shows the ratio of the number of potential interference events recorded at L1 (top panel) and L2 (bottom panel) to total number of hypothesis tests performed at each location for the foregoing detection threshold. A high ratio of potential interference events is reported for both L1 and L2 near Syria (marked with a red dot). This is consistent with the discovery of a strong GNSS jamming source at the Khmeimim Air Base in Syria [106]. Note that the interference "hotspot" appears to the east of the source because the ISS orbit is prograde and the FOTON antenna points in the anti-velocity direction. In other words, the FOTON antenna is exposed to interference only after the ISS passes eastward over an emitter's location.

The high values of the statistic for both L1 and L2 east of Syria indicate that the interference activity in Syria has been persistent over nearly the full interval considered in this chapter, from March 2017 to June 2020. A monthly analysis (not shown) revealed that the source has been transmitting at L2 since at latest March 2017. It began transmitting weak interference at L1 during the second half of 2017, then much stronger interference at L1 during the first quarter of 2018. The interference at L1 and L2 was ongoing in June 2020.

A weaker hotspot is present to the west of the Syrian interference. This hotspot, which emerged in the second half of 2019, is consistent with reports of GNSS interference in the Libyan region [159]. The magenta dots in Fig. 3.8 denote the approximate location of the area in which interference has been documented (33° N, 14° E). Fig. 3.8 also reveals strong L2 interference over mainland China. This interference has been present since at latest March 2017 and was ongoing in June 2020. The green dot in Fig. 3.8, marked at (32° N, 114° E), indicates a hypothesized interference source location based on the shape and location of the observed hotspot.

Note that the above method of counting potential interference events based on CINR degradation ignores cases where interference might lead to complete loss of track of some or all GPS signals. However, the data from the ISS shows that FOTON does not lose track of authentic GNSS signals even when flying by the strong interference source in Syria. In fact, the reported CINR over Syria is well above the weakest signal that FOTON is capable of tracking. As a result, it was concluded that in cases where FOTON seems to

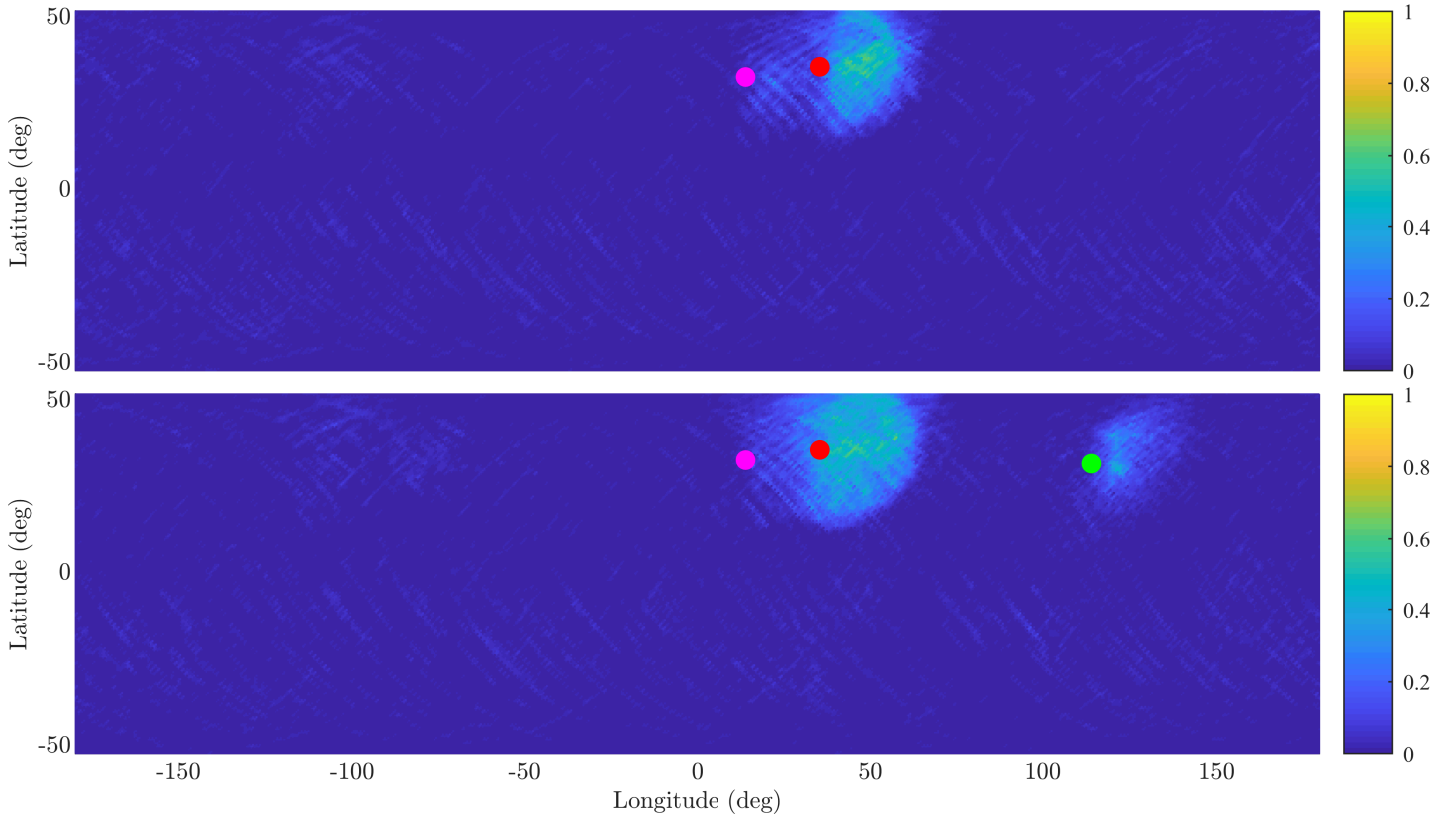


Figure 3.8: Ratio of number of potential GPS L1 (top panel) and L2 (bottom panel) interference events recorded to total number of hypothesis tests performed at each location on the map for the full span of data considered in this chapter, from March 2017 to June 2020. The red dot indicates the reported origin of the Syrian interference in [106]. Another hotspot of interference is apparent to the west of the Syrian interference. The magenta dots denote the approximate location of GNSS interference reports in the Libyan region [159]. In addition to the interference over the Syrian and Libyan regions, strong L2 interference over mainland China is observed. The green dot at (32° N, 114° E) indicates a hypothesized interference source location based on the shape and location of the observed hotspot.

track few or no GPS signals, it is likely due to some abnormal behavior of the receiver, and not due to a potential interference event.

In addition to the global average analysis summarized in Fig. 3.8, it is instructive to examine the time history of receiver reported CINR as the ISS passes over an interference hotspot. Fig. 3.9 shows two such histories for signals within the admissible range of z_r as the ISS goes over the strong interference regions in Syria (Fig. 3.9(a)) and China (Fig. 3.9(b)). Green and blue data points represent range-compensated CINR values for authentic L1 and L2 GNSS signals, respectively, above the applicable threshold, which depends on i, f, and z_r. Light red data points are the same data when below the applicable threshold. Both L1 and L2 signals are declared under interference in Fig. 3.9(a), whereas only L2 signals are declared under interference in Fig. 3.9(b). The brief dip in Fig. 3.9(b) prior to the major dip over China is caused by the Syrian interference. Gaps in the time histories indicate periods with no tracked signals in the admissible zenith angle window.

3.5 Conclusions

Low-earth-orbiting instruments capable of receiving signals in GNSS bands are a powerful tool for detecting GNSS interference emanating from terrestrial sources. Data from one such instrument, the FOTON software-defined GNSS receiver, which has been operational on the International Space Station since February 2017, reveal interesting patterns of GNSS interference from March 2017 to June 2020. Previously-reported powerful interference sources

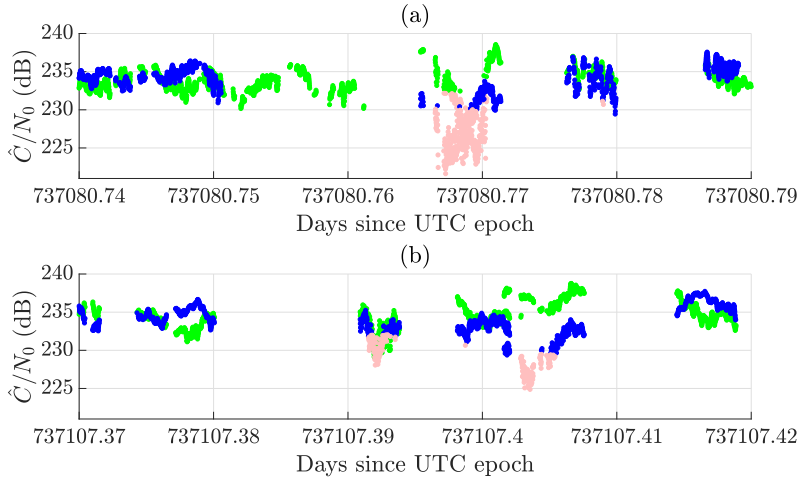


Figure 3.9: Time histories of range-compensated receiver-reported CINR as the ISS flies over potential GPS interference zones over Syria and China.

active in Syria and Libya were confirmed with the FOTON-reported C/N_0 data. In addition, this study reports the approximate onset of these interference activities, as well as their persistent and ongoing nature. Remarkably, the global analysis revealed another previously-unreported interference hotspot in mainland China that only affects the GPS L2 frequency band.

Part III

Towards Low-Cost Robust PNT

Chapter 4

All-Weather sub-50-cm Radar-Inertial Positioning

4.1 Abstract

Deployment of automated ground vehicles beyond the confines of sunny and dry climes will require sub-lane-level positioning techniques based on radio waves rather than near-visible-light radiation. Like human sight, lidar and cameras perform poorly in low-visibility conditions. This chapter develops and demonstrates a novel technique for robust sub-50-cm-accurate urban ground vehicle positioning based on all-weather sensors. The technique incorporates a computationally-efficient globally-optimal radar scan batch registration algorithm into a larger estimation pipeline that fuses data from commercially-available low-cost automotive radars, low-cost inertial sensors, vehicle motion constraints, and, when available, precise GNSS measurements. Performance

---

This chapter is based on:

Lakshay Narula, Peter A Iannucci, and Todd E Humphreys. All-weather sub-50-cm radar-inertial positioning. *Field Robotics*, 2020. Submitted for review.

Lakshay Narula, Peter A Iannucci, and Todd E Humphreys. Automotive-radar-based 50-cm urban positioning. In *Proceedings of the IEEE/ION PLANSx Meeting*, 2020.

is evaluated on an extensive and realistic urban data set. Comparison against ground truth shows that during 60 min of GNSS-denied driving in the urban center of Austin, TX, the technique maintains 95$^{\text{th}}$-percentile errors below 50 cm in horizontal position and 0.5° in heading.

4.2 Introduction

Development of automated ground vehicles (AGVs) has spurred research in lane-keeping assist systems, automated intersection management [53], tight-formation platooning, and cooperative sensing [37, 77], all of which demand accurate (e.g., 50-cm at 95%) ground vehicle positioning in an urban environment. But the majority of positioning techniques developed thus far depend on lidar or cameras, which perform poorly in low-visibility conditions such as snowy whiteout, dense fog, or heavy rain. Adoption of AGVs in many parts of the world will require all-weather localization techniques.

Radio-wave-based sensing techniques such as radar and GNSS (global navigation satellite system) remain operable even in extreme weather conditions [173] because their longer-wavelength electromagnetic radiation penetrates snow, fog, and rain. Carrier-phase-differential GNSS (CDGNSS) has been successfully applied for the past two decades as an all-weather decimeter-accurate localization technique in open-sky conditions. Proprioceptive sensors such as IMUs also continue to operate regardless of external conditions. Coupling a CDGNSS receiver with a tactical-grade inertial sensor, as in [72, 119, 133, 176] delivers robust high-accuracy positioning even during the extended

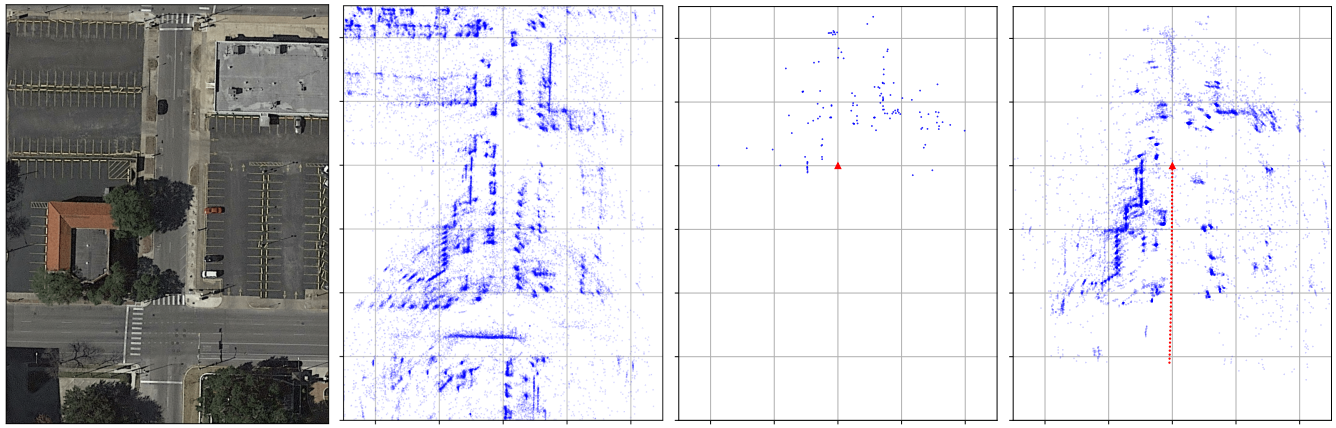
signal outages common in the urban environment, but such systems are far too expensive for widespread deployment on AGVs. Recent work has shown that 20-cm-accurate (95%) CDGNSS positioning is possible at low cost even in dense urban areas, but solution availability remains below 90%, with occasional long gaps between high-accuracy solutions [67]. Moreover, the global trend of increasing radio interference in the GNSS bands, whether accidental or deliberate [65], underscores the need for GNSS-independent localization: GNSS jamming cannot be allowed to paralyze an area's automated vehicle networks.

Clearly, there is a need for AGV localization that is low cost, accurate at the sub-50-cm level, robust to low-visibility conditions, and continuously available. This chapter is the first to establish that low-cost inertial- and automotive-radar-based localization can meet these criteria.

Mass-market commercialization has brought the cost of automotive radar down enough that virtually all current production vehicles are equipped with at least one radar unit, which serves as the primary sensor for adaptive cruise control and automatic emergency braking. But use of automotive radar for localization faces the significant challenges of data sparsity and noise: an automotive radar scan has vastly lower resolution than a camera image or a dense lidar scan, and is subject to high rates of false detection (clutter) and missed detection. As such, it is nearly impossible to deduce semantic information or to extract distinctive environmental features from an individual radar scan. This is clear from Fig. 4.1c, which shows a sparse smattering

of reflections from a single composite scan using three radar units. The key to localization is to aggregate sequential scans into a batch, as in Fig. 4.1d, where environmental structure is clearly evident. Even still, the data remain so sparse that localization based on traditional machine vision feature extraction and matching is not promising. Additionally, stable short-term odometry is a pre-requisite for aggregating radar scans, which in itself is a challenge when dealing with low-cost inertial sensing.

This chapter proposes a two-step process for radar-based localization. The first is the mapping step: creation of a geo-referenced two-dimensional aggregated map of all radar targets across an area of interest. Fig. 4.1b shows such a map, hereafter referred to as a radar map. The full radar map used throughout this chapter, of which Fig. 4.1b is a part, was constructed with the benefit of a highly stable inertial platform so that a trustworthy ground truth map would be available against which maps generated by other techniques could be compared. But an expensive inertial system or dedicated mobile mapping vehicle is not required to create a radar map. Instead, it can be crowd-sourced from the very user vehicles that will ultimately exploit the map for localization. During periods of favorable lighting conditions and good visibility, user vehicles can exploit a combination of low-cost CDGNSS, as in [67], and GNSS-aided visual simultaneous localization and mapping, as in [113], to achieve the continuous decimeter-and-sub-degree-accurate geo-referenced position and orientation (pose) required to lay down an accurate radar map. In other words, the radar map can be *created* when visibility is good and then



(a) (b) (c) (d)

Figure 4.1: Panel (a) shows a satellite view of the environment being mapped with automotive radar. Panel (b) shows the generated radar map point cloud with vehicle pose obtained from a reference localization system. Note the repeating structure along the road side due to parked vehicles. An individual radar scan obtained during localization is shown in panel (c), along with the red triangle denoting vehicle location and heading. The scan is sparse and contains significant clutter, making it challenging to register to the prior map. Panel (d) shows a batch of radar scans during localization, with the red dots denoting the vehicle trajectory over the past five seconds. The batch captures the underlying structure which can be registered to the prior map.

exploited at any later time, such as during times of poor visibility.

Despite aggregation over multiple vehicle passes, the sparse and cluttered nature of automotive radar data is evident from the radar map shown in Fig. 4.1b: the generated point cloud is much less dense and has a substantially higher fraction of spurious returns than a typical lidar-derived point cloud, making automotive-radar-based localization a significantly more challenging problem.

The second step of this chapter's technique is the localization step. Using a combination of all-weather odometric techniques such as inertial sensing, radar odometry, and ground vehicle dynamics constraints, a sensor fusion filter continually tracks the changes in vehicle pose over time. Over the latest short interval (e.g., 5 s), pose estimates from the filter are used to spatially organize the multiple radar scans taken over the interval and generate what is hereafter referred to as a batch of scans, or batch for short. Fig. 4.1d shows a five-second batch terminating at the same location as the individual scan in Fig. 4.1c. In contrast to the individual scan, some environmental structure emerges in the batch of scans, making robust registration to the map feasible. Even so, the localization problem remains challenging due to the dynamic radar environment: note the absence of parked cars on the left side of the street during localization. The batch of scans is matched against the prior map of the surroundings to estimate the pose offset of the batch from the truth. This pose offset is then applied as a measurement to the sensor fusion filter to correct odometric drift.

Contributions. This chapter's overall contribution is a robust pipeline for all-weather sub-50-cm urban ground vehicle positioning. This pipeline incorporates a computationally-efficient correlation-maximization-based globally-optimal radar scan registration algorithm that estimates a two-dimensional translational and a one-dimensional rotational offset between a prior radar map and a batch of current scans. Significantly, the registration algorithm can be applied to the highly sparse and cluttered data produced by commercially-available low-cost automotive radars. Maximization of correlation is shown to be equivalent to minimization of the L^2 distance between the prior map and the batch probability hypothesis densities. The pipeline supports the construction of the radar registration estimate and optimally fuses it with inertial measurements, radar range rate measurements, ground vehicle dynamics constraints, and cm-accurate GNSS measurements, when available. A novel technique for online estimation of the vehicle center of rotation is introduced, and calibration of various other extrinsic parameters necessary for optimal sensor fusion is described.

This chapter also presents a thorough evaluation of the positioning pipeline on the large-scale dataset described in [111]. Data from automotive sensors are collected over two 1.5 h driving sessions through the urban center of Austin, TX on two separate days specifically chosen to provide variety in traffic and parking patterns. Comparison with a post-processed ground truth trajectory shows that proposed pipeline maintains 95^{th}-percentile errors below 35 cm in horizontal position and 0.5° in heading during 60 min of GNSS-denied

driving.

Organization of the rest of this chapter. Sec. 4.3 establishes the significance of this contribution in view of the prior work in related fields. The radar batch-based pose estimation technique for the low-cost automotive radar sensor model is developed in Sec. 4.4. Sec. 4.5 describes the overall sensor fusion architecture involving inertial sensing, GNSS, motion constraints, and radar measurements. Implementation details and experimental results from field evaluation are presented in Sec. 4.6, and Sec. 4.7 provides concluding remarks.

4.3 Related Work

This section reviews a wide variety of literature on mapping and localization with radar and radar-inertial sensing. This includes prior work on point cloud alignment and image registration techniques, occupancy grid-based mapping and localization, random-finite-set-based mapping and localization, and inertial-aided mapping and localization.

Related work in point cloud alignment. A radar-based map can have many different representations. One obvious representation is to store all the radar measurements as a point cloud. Fig. 4.1b is an example of this representation. Localization within this map can be performed with point cloud registration techniques like the iterative closest point (ICP) algorithm. ICP is known to converge to local minima which may occur due to outlying points that do not have correspondences in the two point clouds being aligned.

A number of variations and generalizations of ICP robust to such outliers have been proposed in the literature [35,56,62,70,107,156,163]. A few of these have been applied specifically to automotive radar data [62,163]. But the technique in [163] is only evaluated on a 5 min dataset, while [62] performs poorly on datasets larger than 1 km.

This chapter steers away from ICP and its gradient-based variants because automotive radar data in urban areas exhibit another source of incorrect-but-plausible registration solutions which are not addressed in the above literature: repetitive structure, e.g., due to a series of parked cars, may result in multiple locally-optimal solutions within 2–3 m of the globally-optimal solution. Gradient-based techniques which iteratively estimate correspondences based on the distance between pairs of points are susceptible to converge to such locally-optimal solutions. Accordingly, the batch-based pose estimator proposed in this chapter is designed to approximate the globally-optimal solution.

In contrast to ICP and its variants, globally-optimal point cloud registration can be achieved by performing global point correspondence based on distinctive feature descriptors [12, 33, 34]. All of these works use a sophisticated mechanically-rotating radar unit that is not expected to be available on an AGV. Feature description and matching on the low-cost automotive radars used in this chapter is likely to be fragile. Even when using the mechanically-rotating radar, [13] shows that a correlation-based approach, such as the one developed in this chapter, outperforms other feature-descriptor-based point

cloud methods.

Related work in image registration and occupancy grid techniques. Occupancy grid mapping and localization techniques have been traditionally applied for lidar-based systems, and recent work in [135, 136] has explored similar techniques with automotive radar data. In contrast to batch-based pose estimation described in this chapter, both [136] and [135] perform particle-filter based localization with individual scans, as is typical for lidar-based systems. These methods were only evaluated on small-scale datasets collected in a parking lot, and even so, the reported meter-level localization accuracy is not sufficient for lane-level positioning.

Occupancy grid maps are similar to camera-based top-down images, and thus may be aligned with image registration techniques, that may be visual-descriptor-based [32, 63] or correlation-based [174]. Reliable extraction and matching of visual features, e.g., SIFT or SURF, is significantly more challenging with automotive radar data. Correlation-based registration is more robust [13, 174], and forms the basis of one of the components in this chapter. In contrast to prior work [13, 174], this chapter provides a probabilistic interpretation for the correlation operation. The mechanically-rotating radar of [13] allows correlation-based pose estimation based on a single scan of radar data. But for the low-cost automotive radars used in this chapter, it becomes necessary to accumulate radar scans over time, which requires integration with other odometric sensors. This chapter develops and demonstrates a complete sensor fusion pipeline around radar-based pose estimation and evaluates its

performance on a large urban dataset.

Related work in random finite set mapping and localization. The occupancy grid representation commonly used in robotics is an approximation to the probability hypothesis density (PHD) function [51,90]: a concept first introduced in the random finite set (RFS) based target tracking literature. Unsurprisingly, PHD- and RFS-based mapping and localization have been previously studied in [47, 101, 152]. In contrast to occupancy grid-based methods, techniques in [47, 101, 152] make the point target assumption where no target may generate more than one measurement in a single scan, and no target may occlude another target. However, in reality, planar and extended targets such as walls and building fronts are commonplace in the urban AGV environment. Mapping of ellipsoidal extended targets has recently been proposed in [54], but occlusions are not modeled and only simulation results are presented.

Related work in inertial-aided mapping and localization. This chapter couples radar batch-based pose estimation with other all-weather automotive sensors such as IMU and GNSS. Inertial aiding has been widely applied in visual- and lidar-based mapping and localization [36,55,86,104,125,149,172]. This chapter extends inertial-aiding to sensors that can operate under harsh weather conditions. Recently, radar measurements have been applied to constrain IMU position drift in [14]. Radar-inertial odometry for indoor robots has been proposed in [5, 75]. This chapter is the first to integrate low-cost automotive radars with inertial sensing, GNSS, and ground vehicle dynamics

for lane-level accurate positioning in challenging urban environments.

4.4 Radar-Batch-Based Pose Estimation

This section describes the formulation of the radar-batch-based pose estimation method introduced in this chapter. It first details the statistical motivation behind the method, and then develops an efficient approximation to the globally-optimal estimator. The output of this estimator acts as one of the measurements provided to the overall localization system presented later in Sec. 4.5.

4.4.1 Pose Estimation using Probability Hypothesis Density

For the purpose of radar-based pose estimation, an AGVs environment can be described as a collection of arbitrarily shaped radar reflectors in a specific spatial arrangement. These radar reflectors exist in a two-dimensional plane since the automotive radars used in this work only provide range and azimuth measurements. Assuming sufficient temporal permanence of this environment, radar-equipped AGVs make sample measurements of the underlying structure over time. This section assumes a static radar environment, dynamic objects in the environment are excluded in a pre-processing step (Sec. 4.5.4.3).

4.4.1.1 The Probability Hypothesis Density Function

A probabilistic description of the radar environment is required to set up radar-based pose estimation as an optimization problem. This chapter

chooses the PHD function [90] representation of the radar environment. The PHD at a given location gives the density of the expected number of radar reflectors at that location. For a static radar environment, the PHD $D(\boldsymbol{x})$ at a location $\boldsymbol{x} \in \mathcal{X}$ can be written as

$$D(\boldsymbol{x}) = I \cdot p(\boldsymbol{x})$$

where $\mathcal{X}$ is the set of all locations in the environment, $p(\boldsymbol{x})$ is a probability density function such that $\int p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = 1$, and I, the intensity, is the total number of radar reflectors in the environment. For a time-varying radar environment, both I and $p(\boldsymbol{x})$ are functions of time. For $\mathcal{A} \subset \mathcal{X}$, the expected number of radar reflectors in $\mathcal{A}$ is given as

$$I_{\mathcal{A}} = \int_{\mathcal{A}} D(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$$

4.4.1.2 Estimating Vehicle State from PHDs

Let $D_{\mathrm{m}}(\boldsymbol{x})$ denote the "map" PHD function representing the distribution of radar reflectors in an environment, estimated as a result of mapping with known vehicle poses. During localization, the vehicle makes a radar scan, or a series of consecutive radar scans. A natural solution to the pose estimation problem may be stated as the vehicle pose which maximizes the likelihood of the observed batch of scans, given that the scan was drawn from $D_{\mathrm{m}}(\boldsymbol{x})$ [107]. This maximum likelihood estimate (MLE) has many desirable properties such as asymptotic efficiency. However, the MLE solution is known to be sensitive to outliers that may occur if the batch of scans was sampled from a slightly dif-

ferent PHD, e.g., due to variations in the radar environment between mapping and localization [70].

A more robust solution to the PHD-based pose estimation problem may be stated as follows. Let $\boldsymbol{\Theta}$ denote the vector of parameters of the rigid or non-rigid transformation $\mathcal{T}$ between the vehicle's prior belief of its pose, and its true pose. For example, in case of a two-dimensional rigid transformation (for linear phased-array radars), $\boldsymbol{\Theta} = [\Delta x, \Delta y, \Delta\phi]^\top$, where Δx and Δy denote a two-dimensional position and $\Delta\phi$ denotes heading. Also, let $D_\mathrm{b}(\boldsymbol{x}')$ denote a local "batch" PHD function estimated from a batch of scans during localization, defined over $\boldsymbol{x}' \in \mathcal{A} \subset \mathcal{X}$. This PHD is represented in the coordinate system consistent with the vehicle's prior belief, such that $\boldsymbol{x}' = \mathcal{T}_{\boldsymbol{\Theta}}(\boldsymbol{x})$. Estimating the vehicle pose during localization is defined as estimating $\boldsymbol{\Theta}$ such that some distance metric between the PHDs $D_\mathrm{m}(\boldsymbol{x})$ and $D_\mathrm{b}(\boldsymbol{x}')$ is minimized.

This chapter chooses the L^2 distance between $D_\mathrm{m}(\boldsymbol{x})$ and $D_\mathrm{f}(\boldsymbol{x}')$ as the distance metric to be minimized. As compared to the MLE which minimizes Kullback-Leibler divergence, L^2 minimization trades off asymptotic efficiency for robustness to measurement model inaccuracy [70]. The L^2 distance $d_{L^2}(\boldsymbol{\Theta})$ to be minimized is given as

$$d_{L^2}(\boldsymbol{\Theta}) = \int_{\mathcal{A}} \left(D_\mathrm{m}(\boldsymbol{x}) - D_\mathrm{b}(\mathcal{T}_{\boldsymbol{\Theta}}(\boldsymbol{x}))\right)^2 \mathrm{d}\boldsymbol{x}$$

For rigid two-dimensional transformations, it can be shown as follows that minimizing the L^2 distance between the PHDs is equivalent to maximiza-

tion of the cross-correlation between the PHDs.

$$\begin{aligned}\widehat{\boldsymbol{\Theta}} &= \underset{\boldsymbol{\Theta}'}{\operatorname{argmin}} \int_{\mathcal{A}} \left(D_{\mathrm{m}}(\boldsymbol{x}) - D_{\mathrm{b}}(\mathcal{T}_{\boldsymbol{\Theta}'}(\boldsymbol{x}))\right)^2 \mathrm{d}\boldsymbol{x} \\ &= \underset{\boldsymbol{\Theta}'}{\operatorname{argmin}} \left[\int_{\mathcal{A}} D_{\mathrm{m}}^2(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \int_{\mathcal{A}} D_{\mathrm{b}}^2(\mathcal{T}_{\boldsymbol{\Theta}'}(\boldsymbol{x}))\mathrm{d}\boldsymbol{x} - 2\int_{\mathcal{A}} D_{\mathrm{m}}(\boldsymbol{x}) D_{\mathrm{b}}(\mathcal{T}_{\boldsymbol{\Theta}'}(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}\right]\end{aligned}$$

Note that the first term above is fixed during optimization, while the second term is invariant under rigid transformation. As a result, the above optimization is equivalent to maximizing the cross-correlation:

$$\widehat{\boldsymbol{\Theta}} = \underset{\boldsymbol{\Theta}'}{\operatorname{argmax}} \int_{\mathcal{A}} D_{\mathrm{m}}(\boldsymbol{x}) D_{\mathrm{b}}(\mathcal{T}_{\boldsymbol{\Theta}'}(\boldsymbol{x}))\mathrm{d}\boldsymbol{x} \tag{4.1}$$

For differentiable D_{m} and D_{b}, the above optimization can be solved with gradient-based methods. However, the cross-correlation maximization problem in the urban AGV environment may have locally optimal solutions in the vicinity of the global minimum due to repetitive structure of radar reflectors. In applications with high integrity requirements, a search for the globally optimal solution is necessary. This chapter notes that if the PHDs in (4.1) were to be discretized in $\boldsymbol{x}$, then the cross-correlation values can be evaluated exhaustively with computationally efficient techniques. Let $\boldsymbol{x}_{pq}$ denote the location at the (p, q) translational offset in discretized $\mathcal{A}$. Then

$$\widehat{\boldsymbol{\Theta}} = \underset{\boldsymbol{\Theta}'}{\operatorname{argmax}} \sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} D_{\mathrm{m}}(\boldsymbol{x}_{pq}) D_{\mathrm{b}}(\lfloor \mathcal{T}_{\boldsymbol{\Theta}'}(\boldsymbol{x}_{pq}) \rceil) \tag{4.2}$$

where $\lfloor . \rceil$ denotes the nearest grid point in the discretized space.

The technique developed above relies on the PHDs D_{m} and D_{b}. The next subsections detail the recipe for estimating these PHDs from the radar observations.

4.4.2 Estimating the map PHD from measurements

This section addresses the procedure to estimate the map PHD $D_{\mathrm{m}}(\boldsymbol{x})$ from radar measurements. This chapter works with an occupancy grid map (OGM) approximation to the continuous PHD function. In [51], it has been shown that the PHD representation is a limiting case of the OGM as the grid cell size becomes vanishingly small. Intuitively, let c_{pq} denote the grid cell region with center $\boldsymbol{x}_{pq}$, and let δc_{pq} denote the area of this grid cell, which is small enough such that no more than one reflector may be found in any cell. Let $p_{pq}(O)$ denote the occupancy probability of c_{pq}, and let $\mathcal{A}$ be defined as the region formed by the union of all c_{pq} whose centers $\boldsymbol{x}_{pq}$ fall within $\mathcal{A}$. Then, the expected number of radar reflectors $\mathbb{E}[|\mathcal{A}|]$ in $\mathcal{A}$ is given by

$$\begin{aligned}\mathbb{E}[|\mathcal{A}|] = \sum_{c_{pq}\in\mathcal{A}} p_{pq}(O) &= \sum_{c_{pq}\in\mathcal{A}} \frac{p_{pq}(O)}{\delta c_{pq}}\delta c_{pq} \\ &\triangleq \sum_{c_{pq}\in\mathcal{A}} \bar{D}(\boldsymbol{x}_{pq})\delta c_{pq} \\ &= \int_{\mathcal{A}} \bar{D}(\boldsymbol{x}_{pq})\mathrm{d}\boldsymbol{x}, \quad \text{as} \lim_{\delta c_{pq}\to 0}\end{aligned}$$

where $\bar{D}(\boldsymbol{x}_{pq}) \equiv \frac{p_{pq}(O)}{\delta c_{pq}}$ can be considered to be an approximation of the PHD $D(\boldsymbol{x})$ for $\boldsymbol{x} \in c_{pq}$ since its integral over $\mathcal{A}$ is equal to the expected number of reflectors in $\mathcal{A}$.

The advantage of working with an OGM approximation of the PHD is two-fold: first, since the OGM does not attempt to model individual objects, it is straightforward to represent arbitrarily-shaped objects, and second, in contrast to the "point target" measurement model assumption in standard PHD

filtering, the OGM can straightforwardly model occlusions due to extended objects.

At this point, the task of estimating $D_{\mathrm{m}}(\boldsymbol{x})$ has been reduced to estimating the occupancy probability of each grid cell in discretized $\mathcal{A}$. Each grid cell c_{pq} takes up one of two states: occupied (O) or free (F). Based on the radar measurement $\boldsymbol{z}_k$ at each time k, the Bernoulli probability distribution of such binary state cells may be recursively updated with the binary Bayes filter. In particular, let $\boldsymbol{z}_{1:k}$ denote all radar measurements made up to time k, and let

$$l_{pq}^{k}(O) \equiv \log \frac{p_{pq}(O \mid \boldsymbol{z}_{1:k})}{1 - p_{pq}(O \mid \boldsymbol{z}_{1:k})} \tag{4.3}$$

denote the *log odds ratio* of c_{pq} being in state O. Also define $l_{pq}^{0}(O)$ as

$$l_{pq}^{0}(O) \equiv \log \frac{p_{pq}(O)}{1 - p_{pq}(O)}$$

with $p_{pq}(O)$ being the prior belief on the occupancy state of c_{pq} before any measurements are made. With these definitions, the binary Bayes filter update is given by [153]

$$l_{pq}^{k}(O) = \log \frac{p_{pq}(O \mid \boldsymbol{z}_k)}{1 - p_{pq}(O \mid \boldsymbol{z}_k)} - l_{pq}^{0}(O) + l_{pq}^{k-1}(O) \tag{4.4}$$

where $p_{pq}(O \mid \boldsymbol{z}_k)$ is known as the *inverse* sensor model: it describes the probability of c_{pq} being in state O, given only the latest radar scan $\boldsymbol{z}_k$.

The required occupancy probability $p_{pq}(O \mid \boldsymbol{z}_{1:k})$ is easy to compute from the log odds ratio in (4.3). Observe that the inverse sensor model

$p_{pq}(O \mid \boldsymbol{z}_k)$, in addition to the prior occupancy belief $p_{pq}(O)$, completely describes the procedure for estimating the OGM from radar measurements, and hence approximating the PHD. Adapting $p_{pq}(O \mid \boldsymbol{z}_k)$ to the characteristics of the automotive radar sensors, however, is not straightforward, and is discussed next.

4.4.3 Automotive Radar Inverse Sensor Model

This section addresses the challenge of adapting the inverse sensor model $p_{pq}(O \mid \boldsymbol{z}_k)$ to the measurement characteristics of automotive radar sensors. Fig. 4.2 shows a simplified radar scan $\boldsymbol{z}_k$ of an underlying occupancy grid. For clarity of exposition, four distinct categories of grid cells in Fig. 4.2 are defined below:

- *Type A*: Grid cells in the vicinity of a radar range-azimuth return.
- *Type B*: Grid cells along the path between the radar sensor and *Type A* grid cells.
- *Type C*: Grid cells in the "viewshed" of the radar sensor, i.e., within the radar field-of-view and not shadowed by another object, but not of *Type A* or *Type B*.
- *Type D*: Grid cells outside the field-of-view of the radar (*Type D1*) or shadowed by other objects closer to the radar (*Type D2*).

The inverse sensor model must choose a $p_{pq}(O \mid \boldsymbol{z}_k)$ value for each of these types of grid cells. In the following, the subscript pq is dropped for cleaner

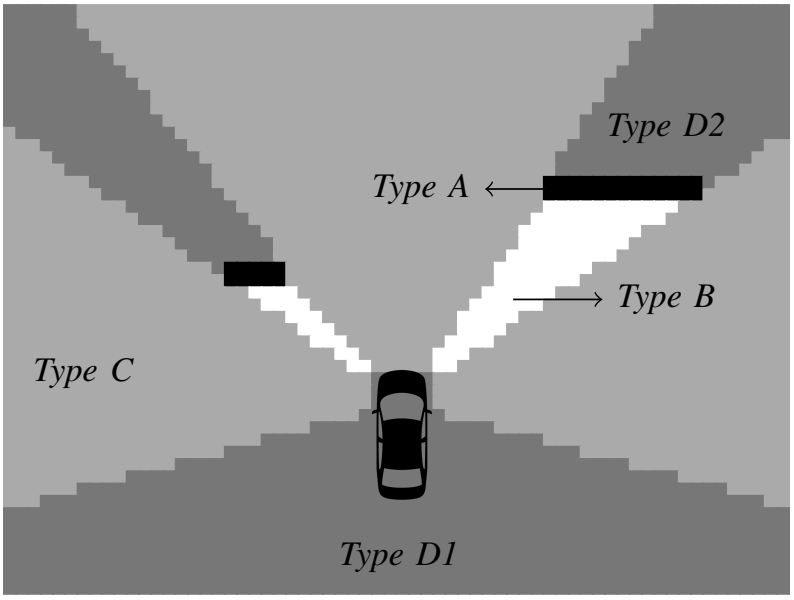


Figure 4.2: Schematic diagram showing four types of grid cells considered in the inverse sensor model.

notation.

4.4.3.1 Conventional Choices

Since $\boldsymbol{z}_k$ provides no additional information on *Type D* grid cells, the occupancy in these cells is conditionally independent of $\boldsymbol{z}_k$, that is

$$p^D(O \mid \boldsymbol{z}_k) = p(O)$$

where $p(O)$ is the prior probability of occupancy defined earlier in Sec. 4.4.1.

Grid cells of *Type B* and *Type C* may be hypothesized to have low occupancy probability, since these grid cells were scanned by the sensor but

no return was obtained. As a result, conventionally

$$p^B(O \mid \boldsymbol{z}_k) \leq p(O)$$

and

$$p^C(O \mid \boldsymbol{z}_k) \leq p(O)$$

Finally, grid cells of *Type A* may be hypothesized to have higher occupancy probability, since a return has been observed in the vicinity of these cells. Conventionally,

$$p^A(O \mid \boldsymbol{z}_k) \geq p(O)$$

In the limit, if the OGM grid cell size is comparable to the sensor range and angle uncertainty, or if the number of scans is large enough such that the uncertainty is captured empirically, only the grid cells that contain the sensor measurement may be considered to be of *Type A*.

4.4.3.2 Automotive Radar Sensor Characteristics

Intense clutter properties and sparsity of the automotive radar data complicate the choice of the inverse sensor model.

Sparsity. First, sparsity of the radar scan implies that many occupied *Type A* grid cells in the radar environment might be incorrectly categorized as free *Type C* cells. This can be observed in Fig. 4.1. As evidenced by the batch of scans in Fig. 4.1d, the radar environment is "dense" in that many grid cells contain radar reflectors. However, any individual radar scan, such as the one shown in Fig. 4.1c, suggests a much more sparse radar environment. As a

result, a grid cell which is occupied in truth will be incorrectly categorized as *Type C* in many scans, and correctly as *Type A* in a few scans.

The sparsity of radar returns also makes it challenging to distinguish *Type C* cells from cells of *Type D2*. Since many occluding obstacles are not detected in each scan, the occluded cells of *Type D2* are conflated with free cells of *Type C*.

In context of the inverse sensor model, as the radar scan becomes more sparse

$$p^C(O \mid \boldsymbol{z}_k) \to p^D(O \mid \boldsymbol{z}_k)^-$$

where the superscript $-$ denotes a limit approaching from below. Intuitively, approaching $p^D(O \mid \boldsymbol{z}_k)$ implies that the measurement $\boldsymbol{z}_k$ is very sparse in comparison to the true occupancy, and thus does not provide much information on lack of occupancy.

Clutter. Second, there is the matter of clutter. The grid cells in the vicinity of a clutter measurement may be incorrectly categorized as *Type A*, and the grid cells along the path between the radar and clutter measurement may be incorrectly categorized as *Type B*.

In context of the inverse sensor model, as the radar scan becomes more cluttered

$$p^B(O \mid \boldsymbol{z}_k) \to p^D(O \mid \boldsymbol{z}_k)^-$$
$$p^A(O \mid \boldsymbol{z}_k) \to p^D(O \mid \boldsymbol{z}_k)^+$$

where the superscript $+$ denotes a limit approaching from above.

4.4.3.3 A Pessimistic Inverse Sensor Model

The results presented in Sec. 4.6 are based on a pessimistic sensor model, such that $p^B(O \mid \boldsymbol{z}_k) = p^C(O \mid \boldsymbol{z}_k) = p^D(O \mid \boldsymbol{z}_k)$. This model assumes that the radar measurements provide no information about free space in the radar environment.

In particular, the inverse sensor model assumes

$$p^B(O \mid \boldsymbol{z}_k) = p^C(O \mid \boldsymbol{z}_k) = p^D(O \mid \boldsymbol{z}_k) = p(O) = 0.1$$

and

$$p^A(O \mid \boldsymbol{z}_k) = 0.2$$

4.4.4 Estimating the batch PHD from measurements

The procedure for generating an approximation to $D_{\mathrm{b}}(\boldsymbol{x}')$ from a batch of radar measurements is identical to the procedure for generating $D_{\mathrm{m}}(\boldsymbol{x})$ from mapping vehicle data, except that precise, absolute location and orientation data is not available during localization. Instead, pose estimates from the sensor fusion filter described in Sec. 4.5 are used to estimate the relative locations and orientations of each radar scan in the batch, and the scans are transformed into a common coordinate frame before updating the occupancy state of grid cells.

Once the map and batch PHDs have been approximated from radar measurements, the correlation-maximization technique developed in Sec. 4.4.1 can be applied to obtain the estimate $\widehat{\boldsymbol{\Theta}}$. This estimate is handed back to the

sensor fusion filter as a pose offset measurement to constrain the odometric drift during absence of other sources of absolute localization, e.g., GNSS.

4.4.5 Efficient FFT-Based Implementation

This section notes that efficient globally-optimal procedures exist for maximizing the discretized PHD correlation as defined in (4.2), and outlines two optimizations which further reduce the computational complexity of the problem.

4.4.5.1 FFT-based Cross-Correlation

For two-dimensional vehicle state estimation, the cross-correlation in (4.2) is to be maximized over the three parameters of two-dimensional rigid transformation $[\Delta x, \Delta y, \Delta\phi]^\top$.

For a given value of $\Delta\phi$, the cross-correlation can be maximized efficiently over $\Delta\boldsymbol{t} = [\Delta x, \Delta y]^\top$ with FFT-based cross-correlation. The size of the discretized map and batch PHDs to be correlated, denoted $P \times Q$ in (4.2), is limited by the area scanned by the radar over a batch. Without loss of generality, let $P = Q = n$. Due to the convolution property of the FFT, the circular cross-correlation between $n \times n$ matrices $D_{\mathrm{m}}(\boldsymbol{x})$ and $D_{\mathrm{b}}(\mathcal{T}_{\boldsymbol{\Theta}}(\boldsymbol{x}))$ can be computed as

$$D_{\mathrm{m}} * D_{\mathrm{b}} = \mathcal{F}^{-1}\{\mathcal{F}\{D_{\mathrm{m}}(\boldsymbol{x})\} \circ \mathcal{F}\{D_{\mathrm{b}}(-\mathcal{T}_{\boldsymbol{\Theta}}(\boldsymbol{x}))\}\}$$

where $\mathcal{F}$ denotes the FFT operator and $\circ$ the denotes element-wise multiplication operator. To compute the required linear cross-correlation, however,

both D_{m} and D_{b} must be padded with $n/2$ zeros on each side along each dimension, leading to matrices $\check{D}_{\mathrm{m}}$ and $\check{D}_{\mathrm{b}}$ of size $2n \times 2n$. Then, the linear cross-correlation is

$$D_{\mathrm{m}} \star D_{\mathrm{b}} = \mathcal{F}^{-1}\{\mathcal{F}\{\check{D}_{\mathrm{m}}(\boldsymbol{x})\} \circ \mathcal{F}\{\check{D}_{\mathrm{b}}(-\mathcal{T}_{\boldsymbol{\Theta}}(\boldsymbol{x}))\}\} \tag{4.5}$$

The two FFTs and one inverse FFT (IFFT) in (4.5) each have a computational complexity

$$k(2n)^2 \log{(2n)^2} \approx 8kn^2 \log n$$

where k is a constant factor dependent on the FFT implementation. If the number of rotations to be examined are m, this leads to a total complexity of

$$3m \times 8kn^2 \log n = 24kmn^2 \log n \tag{4.6}$$

One observation here is that $\mathcal{F}\{\check{D}_{\mathrm{m}}(\boldsymbol{x})\}$ for the map is independent of $\Delta\phi$, and so must only be computed once. This reduces the overall complexity to $8k(2m+1)n^2 \log n$.

4.4.5.2 Minimal Padding

Typically, the size of the map and batch PHDs to be correlated, given by $P \times Q$, is much larger than the translational offset search space due to initial uncertainty in the vehicle position. In other words, the admissible values of $\Delta\boldsymbol{t}$ lie within a small fraction of the space scanned in the radar batch. Accordingly, the optimization method only requires the linear cross-correlation values within this admissible region. If n_l denotes the size of the

translational search space in discretized PHD coordinates, then D_{m} and D_{b} need only be padded with $n_l/2$ zeros on each side along each dimension, leading to matrices $\check{D}_{\mathrm{m}}$ and $\check{D}_{\mathrm{b}}$ of size $(n+n_l)\times(n+n_l)$. With minimal padding, the overall complexity of FFT-based correlation maximization now reduces to

$$2k(2m+1)(n+n_l)^2\log(n+n_l)\approx 2k(2m+1)n^2\log n$$

where the approximation holds if $n_l \ll n$.

4.4.5.3 The FFT Rotation Theorem

Observe from (4.5) that the method re-computes the FFT after every rotation of the PHD D_{b}. The FFT rotation theorem [127] states that a coordinate rotation in the spatial domain leads to the same coordinate rotation in the frequency domain, that is, if $R_{\Delta\phi}$ represents the rotation matrix which operates on the PHD, then

$$\mathcal{F}\{R_{\Delta\phi}\cdot D_{\mathrm{b}}(\boldsymbol{x})\} = R_{\Delta\phi}\cdot\mathcal{F}\{D_{\mathrm{b}}(\boldsymbol{x})\}$$

This implies that instead of performing m FFTs on rotated replicas of D_{b}, a single FFT may be performed followed by m coordinate rotations. It must be noted that rotation of $\mathcal{F}\{D_{\mathrm{b}}(\boldsymbol{x})\}$ could involve interpolation of values to non-integer indices, which may offset the computational advantage of this method. Experiments conducted as part of this paper suggest that nearest-neighbor interpolation, i.e., assigning value from the nearest integer index (complexity $\mathcal{O}(n^2)$), has no discernible adverse effect on the performance of the algorithm.

With application of the FFT rotation theorem, the overall complexity is reduced to

$$2k(m+2)(n+n_l)^2 \log{(n+n_l)} + m(n+n_l)^2 \approx 2kmn^2 \log n$$

which is a factor of 12 faster than the basic implementation in (4.6).

Algorithm 1 provides the pseudocode for the optimized FFT-based correlation maximization algorithm. For each epoch, the algorithm is provided the prior map point cloud in the true world frame, denoted $\boldsymbol{p}_{\mathrm{m}}^{W}$ and a batch of k radar scans in the body frame, denoted $\boldsymbol{p}_{\mathrm{b},1:k}^{B}$. An initial guess for vehicle position and heading trajectories $\boldsymbol{t}_{1:k}^{V}$ and $\phi_{1:k}^{V}$ is provided in a frame V which is offset from the W frame by a rigid two-dimensional transform parameterized by $\boldsymbol{\Theta} = [\Delta x, \Delta y, \Delta\phi]^{\top}$. The initial guess uncertainties σ_t and σ_ϕ, and the desired discretization steps δt and $\delta\phi$ are also provided. The algorithm must estimate the offset transformation $\widehat{\boldsymbol{\Theta}}$ between W and V.

The `toOGM` routine converts the provided point cloud to an occupancy grid with the desired grid cell size, according to the procedure described in Sec. 4.4.2. The `pad` routine pads the provided array with the desired number of zeros along each dimension on both ends. The three-dimensional matrix `R` holds the linear cross-correlation outputs.

4.5 State Estimation with Sensor Fusion

Thus far, Sec. 4.4 has developed the theory and implementation of the radar batch correlation measurement, which provides an estimate $\widehat{\boldsymbol{\Theta}}$ of the

Algorithm 1: `fastGlobalAlign`

Input : $\boldsymbol{p}_{\mathrm{m}}^{W}$, $\boldsymbol{p}_{\mathrm{b},1:k}^{B}$, $\boldsymbol{t}_{1:k}^{V}$, $\phi_{1:k}^{V}$, σ_t, σ_ϕ, δt, $\delta\phi$

Output: $\widehat{\boldsymbol{\Theta}}$

$D_{\mathrm{m}} = \texttt{toOGM}\big(\boldsymbol{p}_{\mathrm{m}}^{W} - \boldsymbol{t}_{k}^{V}, \delta t\big)$

$\tilde{D}_{\mathrm{m}} = \texttt{FFT2}(\texttt{pad}(D_{\mathrm{m}}, 3\sigma_t))$

$\boldsymbol{p}_{\mathrm{b},1:k}^{V} = R\big(\phi_{1:k}^{V}\big)\boldsymbol{p}_{\mathrm{b},1:k}^{B} + \boldsymbol{t}_{1:k}^{V}$

$D_{\mathrm{b}} = \texttt{toOGM}\Big(\boldsymbol{p}_{\mathrm{b},1:k}^{V} - \boldsymbol{t}_{k}^{V}, \delta t\Big)$

$\tilde{D}_{\mathrm{b}} = \texttt{FFT2}(\texttt{pad}(D_{\mathrm{b}}, 3\sigma_t))$

$n = 3\sigma_t/\delta t$

$m = 3\sigma_\phi/\delta\phi$

for i = *-m:m* **do**

  $\Delta\phi = i\delta\phi$

  $\tilde{D}_{\mathrm{b}}^{\Delta\phi} = \texttt{rotate2}\Big(\tilde{D}_{\mathrm{b}}, \Delta\phi\Big)$

  $\mathsf{R}[i,:,:] = \texttt{IFFT2}\Big(\tilde{D}_{\mathrm{m}} \circ \texttt{conj}\Big(\tilde{D}_{\mathrm{b}}^{\Delta\phi}\Big)\Big)[-n:n, -n:n]$

end

$\widehat{\boldsymbol{\Theta}} = \mathrm{argmax}(\mathsf{R})$

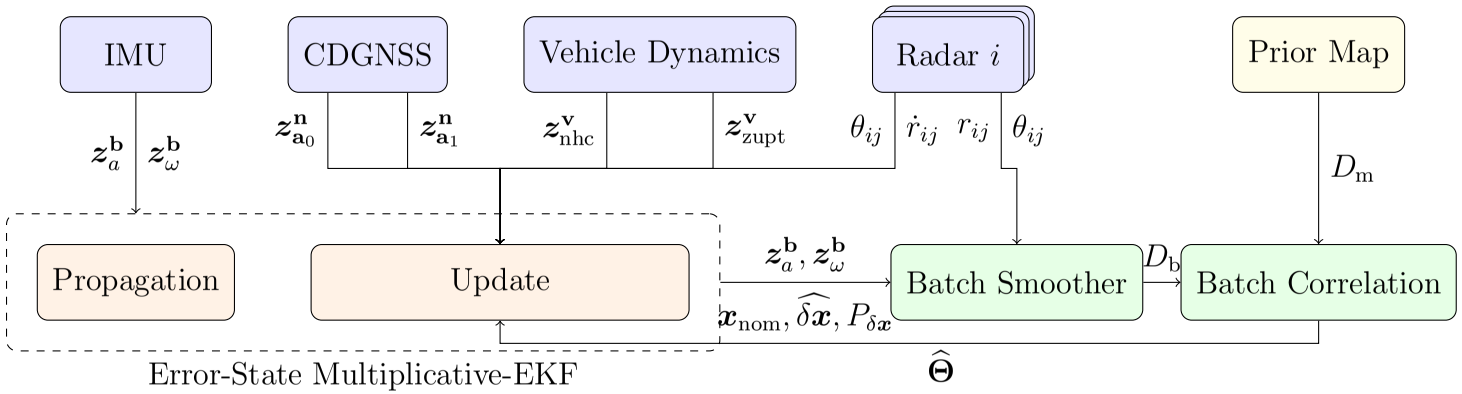


Figure 4.3: Block diagram of the localization pipeline. A low-cost MEMS IMU provides high-rate specific force and angular rate measurements. The error-state multiplicative extended Kalman filter (M-EKF) makes use of cm-accurate CDGNSS position measurements whenever such measurements are available, e.g., in clear-sky GNSS environments. Radial velocity and bearing measurements from low-cost automotive radars are combined with nearly-zero sideslip and vertical speed constraints of a ground vehicle to continually track and limit the errors in inertial navigation. Smoothed batches of radar scans are correlated with a prior map to limit odometric position drift during CDGNSS outages.

3-DoF (degrees-of-freedom) pose offset relative to the prior map. This section details a localization pipeline that incorporates the batch measurement update along with an array of other automotive all-weather sensing modalities to track the full 6-DoF vehicle pose trajectory. The high-rate pose estimates from this pipeline are also used to spatially organize individual scans to form the batch of radar scans used in the batch correlation update.

The choice of sensors available for all-weather localization is limited to radio-frequency sensors such as GNSS and automotive radars, and to proprioceptive sensors such as IMUs and wheel encoders. Any additional domain knowledge, such as properties of ground vehicle dynamics, may also be combined with these sensor measurements.

The localization pipeline in this chapter is developed around a low-cost MEMS (micro electro-mechanical system) IMU. Fig. 4.3 shows a block diagram of the overall pipeline. The error-state multiplicative extended Kalman filter (M-EKF) makes use of cm-accurate CDGNSS position measurements whenever such measurements are available, e.g., in clear-sky GNSS environments. Radial velocity and bearing measurements from low-cost automotive radars are combined with nearly-zero sideslip and vertical speed constraints of a ground vehicle to continually track and limit the errors in inertial navigation. Smoothed batches of radar scans are correlated with a prior map to limit odometric position drift during CDGNSS outages. The following subsections outline the formulation of the estimator, the nonlinear state dynamics, the various measurement models, and the necessary calibration procedures.

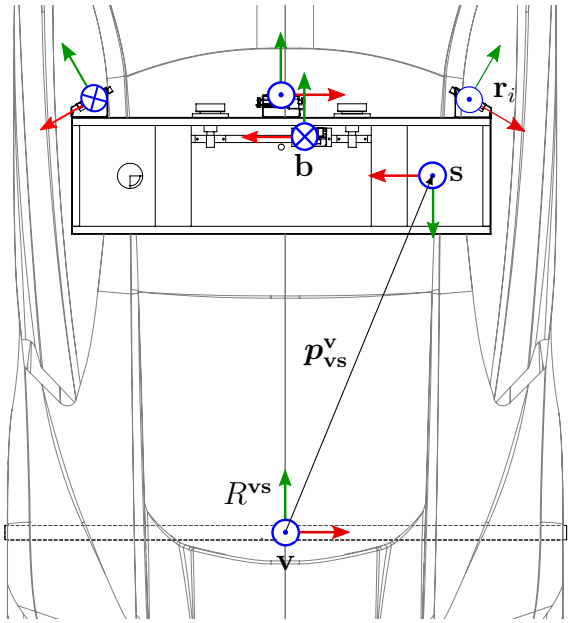


Figure 4.4: The University of Texas Sensorium is an integrated platform for automated and connected vehicle perception research. It includes three automotive radar units, one electronically-scanning radar (ESR) and two short-range radars (SRR2s); stereo visible light cameras; automotive- and industrial-grade IMUs; a dual-antenna, multi-frequency software-defined GNSS receiver; and an internal computer. An iXblue ATLANS-C CDGNSS-disciplined inertial navigation system (not shown) is mounted at the rear of the platform to provide the ground truth trajectory. The vehicle frame $\mathbf{v}$ is located approximately at the center of the line connecting the rear axles.

4.5.1 Sensor Platform & Coordinate Frames

To facilitate the discussion on measurement models and calibration, the sensor-instrumented vehicle and a few related coordinate frames are introduced here. An integrated perception platform called the *Sensorium*, shown schematically in Fig. 4.4, brings together the various low-cost automotive sensors considered in this chapter. Many of these sensors provide measurements

in their respective local frames, leading to a number of different coordinate frames that must be considered.

The IMU *body frame*, denoted $\mathbf{b}$, is the frame defined by the IMU's accelerometer triad.

The *navigation frame*, denoted $\mathbf{n}$, is a local geographical reference frame, e.g., an ENU frame centered at an arbitrary nearby location with the x-axis pointing towards local east, the y-axis pointing towards local north, and the z-axis completing the right-handed coordinate system. The estimator wishes to track the pose trajectory of $\mathbf{b}$ with respect to $\mathbf{n}$.

The *radar frames*, denoted $\mathbf{r}_i$ for the ith radar, are local frames in which the radar sensors report range, range rate, and bearing to a number of targets.

The *vehicle frame*, denoted $\mathbf{v}$, is characterized by the direction in which the vehicle travels when the commanded steering angle is zero. This direction defines the y-axis of $\mathbf{v}$, as shown in Fig. 4.4. The origin of $\mathbf{v}$ is located at the center of rotation of the vehicle.

The *Sensorium frame*, denoted $\mathbf{s}$, is defined by the physical structure of the Sensorium. It is essentially a convenience reference frame in which the nominal lever arm and orientation between different sensors are available per the mechanical specifications of the Sensorium. The origin of $\mathbf{s}$ is arbitrarily chosen to be co-located with one of the GNSS antennas.

4.5.2 Error-State Filtering

The localization system of Fig. 4.3 estimates the following 16-element state vector:

$$\boldsymbol{x}_k = \left[\boldsymbol{p}_k^{\mathbf{n}}, \boldsymbol{v}_k^{\mathbf{n}}, \boldsymbol{q}_k^{\mathbf{nb}}, \boldsymbol{b}_{a,k}^{\mathbf{b}}, \boldsymbol{b}_{\omega,k}^{\mathbf{b}}\right]$$

where $\boldsymbol{p}_k^{\mathbf{n}}$ is the vector from $\mathbf{n}$ to $\mathbf{b}$ at time k expressed in $\mathbf{n}$, $\boldsymbol{v}_k^{\mathbf{n}}$ is the velocity of $\mathbf{b}$ relative to $\mathbf{n}$ at time k expressed in the $\mathbf{n}$ frame, $\boldsymbol{q}_k^{\mathbf{nb}}$ is the quaternion that rotates a vector from $\mathbf{b}$ to $\mathbf{n}$ at time k, and $\boldsymbol{b}_{a,k}^{\mathbf{b}}$ and $\boldsymbol{b}_{\omega,k}^{\mathbf{b}}$ are the accelerometer and gyroscope biases of the IMU at time k, expressed in $\mathbf{b}$.

Note that the vehicle orientation only has three effective degrees-of-freedom since $\boldsymbol{q}_k^{\mathbf{nb}}$ is constrained to be a unit quaternion. Enforcing such a constraint may result in a singular covariance matrix. This issue is typically dealt with an error-state filter [145] where the true state is split into a nominal-state vector

$$\boldsymbol{x}_{\text{nom},k} = \left[\tilde{\boldsymbol{p}}_k^{\mathbf{n}}, \tilde{\boldsymbol{v}}_k^{\mathbf{n}}, \tilde{\boldsymbol{q}}_k^{\mathbf{nb}}, \tilde{\boldsymbol{b}}_{a,k}^{\mathbf{b}}, \tilde{\boldsymbol{b}}_{\omega,k}^{\mathbf{b}}\right]$$

and an error-state vector $\delta\boldsymbol{x}_k$, related by the generalized addition operator $\oplus$ as follows:

$$\boldsymbol{x}_k = \boldsymbol{x}_{\text{nom},k} \oplus \delta\boldsymbol{x}_k$$

where the error-state vector $\delta\boldsymbol{x}_k$ is the minimal 15-element state representation denoted component-wise as follows:

$$\delta\boldsymbol{x}_k = \left[\delta\boldsymbol{p}_k^{\mathbf{n}}, \delta\boldsymbol{v}_k^{\mathbf{n}}, \boldsymbol{\eta}_k^{\mathbf{n}}, \delta\boldsymbol{b}_{a,k}^{\mathbf{b}}, \delta\boldsymbol{b}_{\omega,k}^{\mathbf{b}}\right]$$

The $\oplus$ operator corresponds to usual vector addition for the position, velocity, and bias states. For the orientation state, $\oplus$ is defined as

$$\begin{aligned}\boldsymbol{q}_k^{\mathbf{nb}} &= \tilde{\boldsymbol{q}}_k^{\mathbf{nb}} \oplus \boldsymbol{\eta}_k^{\mathbf{n}} \\ &= \exp_q\left(\frac{\boldsymbol{\eta}_k^{\mathbf{n}}}{2}\right) \odot \tilde{\boldsymbol{q}}_k^{\mathbf{nb}}\end{aligned}$$

where $\exp_q$ denotes the exponential map from $\mathfrak{so}(3)$ to $SO(3)$ [74], represented as a quaternion, and $\odot$ denotes quaternion multiplication. Note that $\boldsymbol{\eta}_k^{\mathbf{n}}$ is parametrized as an orientation deviation in $\mathbf{n}$. A similar formulation may be derived with the orientation deviation expressed in $\mathbf{b}$ [145].

The nonlinear error-state is tracked with an error-state EKF. Owing to the multiplicative orientation dynamics and update, this filter is sometimes referred to as the multiplicative-EKF [43].

4.5.3 State Dynamics

Inertial measurements, collectively denoted $\boldsymbol{u}_k$, are interpreted as control inputs during the state propagation step. The true-state dynamics function $f_k(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k)$ is modeled as

$$\begin{aligned}\boldsymbol{p}_{k+1}^{\mathbf{n}} &= \boldsymbol{p}_k^{\mathbf{n}} + T\boldsymbol{v}_k^{\mathbf{n}} + \frac{T^2}{2}\big(R_k^{\mathbf{nb}}\big(\boldsymbol{z}_{a,k}^{\mathbf{b}} - \boldsymbol{b}_{a,k}^{\mathbf{b}} - \boldsymbol{w}_{a,k}^{\mathbf{b}}\big) + \boldsymbol{g}^{\mathbf{n}}\big) \\ \boldsymbol{v}_{k+1}^{\mathbf{n}} &= \boldsymbol{v}_k^{\mathbf{n}} + T\big(R_k^{\mathbf{nb}}\big(\boldsymbol{z}_{a,k}^{\mathbf{b}} - \boldsymbol{b}_{a,k}^{\mathbf{b}} - \boldsymbol{w}_{a,k}^{\mathbf{b}}\big) + \boldsymbol{g}^{\mathbf{n}}\big) \\ \boldsymbol{q}_{k+1}^{\mathbf{nb}} &= \boldsymbol{q}_k^{\mathbf{nb}} \odot \exp_q\left(\frac{T}{2}\big(\boldsymbol{z}_{\omega,k}^{\mathbf{b}} - \boldsymbol{b}_{\omega,k}^{\mathbf{b}} - R_k^{\mathbf{bn}}\boldsymbol{\omega}_{\mathbf{e}}^{\mathbf{n}} - \boldsymbol{w}_{\omega,k}^{\mathbf{b}}\big)\right) \\ \boldsymbol{b}_{a,k+1}^{\mathbf{b}} &= \boldsymbol{b}_{a,k}^{\mathbf{b}} + \boldsymbol{w}_{\boldsymbol{b}_a,k}^{\mathbf{b}} \\ \boldsymbol{b}_{\omega,k+1}^{\mathbf{b}} &= \boldsymbol{b}_{\omega,k}^{\mathbf{b}} + \boldsymbol{w}_{\boldsymbol{b}_\omega,k}^{\mathbf{b}}\end{aligned}$$

where T is the propagation duration, $R_k^{\mathbf{nb}}$ is the rotation matrix representation of $\boldsymbol{q}_k^{\mathbf{nb}}$, $\boldsymbol{z}_{a,k}^{\mathbf{b}}$ and $\boldsymbol{z}_{\omega,k}^{\mathbf{b}}$ are the IMU specific force and angular rate measurements, respectively, $\boldsymbol{w}_{a,k}$ and $\boldsymbol{w}_{\omega,k}$ are the IMU specific force and angular rate white noise, respectively, $\boldsymbol{g}^{\mathbf{n}} \approx [0, 0, -9.8\,\mathrm{m\,s^{-2}}]$ is the acceleration due to gravity after compensation for the centripetal force due to earth's rotation, and $\boldsymbol{\omega}_{\mathbf{e}}^{\mathbf{n}}$ is the angular rate of the earth with respect to an inertial frame. The accelerometer and gyroscope biases are modeled as random walk processes driven by white noise $\boldsymbol{w}_{\boldsymbol{b}_a,k}^{\mathbf{b}}$ and $\boldsymbol{w}_{\boldsymbol{b}_\omega,k}^{\mathbf{b}}$, respectively, whose variances are derived from the IMU bias instability parameters [168].

The nominal-state dynamics function $f_{\mathrm{nom},k}(\boldsymbol{x}_{\mathrm{nom},k}, \boldsymbol{u}_k, \boldsymbol{w}_k)$ is similar to $f_k(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k)$:

$$
\begin{aligned}
\tilde{\boldsymbol{p}}_{k+1}^{\mathbf{n}} &= \tilde{\boldsymbol{p}}_k^{\mathbf{n}} + T\tilde{\boldsymbol{v}}_k^{\mathbf{n}} + \frac{T^2}{2}\Big(\tilde{R}_k^{\mathbf{nb}}\Big(\boldsymbol{z}_{a,k}^{\mathbf{b}} - \tilde{\boldsymbol{b}}_{a,k}^{\mathbf{b}}\Big) + \boldsymbol{g}^{\mathbf{n}}\Big) \\
\tilde{\boldsymbol{v}}_{k+1}^{\mathbf{n}} &= \tilde{\boldsymbol{v}}_k^{\mathbf{n}} + T\Big(\tilde{R}_k^{\mathbf{nb}}\Big(\boldsymbol{z}_{a,k}^{\mathbf{b}} - \tilde{\boldsymbol{b}}_{a,k}^{\mathbf{b}}\Big) + \boldsymbol{g}^{\mathbf{n}}\Big) \\
\tilde{\boldsymbol{q}}_{k+1}^{\mathbf{nb}} &= \tilde{\boldsymbol{q}}_k^{\mathbf{nb}} \odot \exp_q\left(\frac{T}{2}\Big(\boldsymbol{z}_{\omega,k}^{\mathbf{b}} - \tilde{\boldsymbol{b}}_{\omega,k}^{\mathbf{b}} - \tilde{R}_k^{\mathbf{bn}}\boldsymbol{\omega}_{\mathbf{e}}^{\mathbf{n}}\Big)\right) \\
\tilde{\boldsymbol{b}}_{a,k+1}^{\mathbf{b}} &= \tilde{\boldsymbol{b}}_{a,k}^{\mathbf{b}} \\
\tilde{\boldsymbol{b}}_{\omega,k+1}^{\mathbf{b}} &= \tilde{\boldsymbol{b}}_{\omega,k}^{\mathbf{b}}
\end{aligned}
$$

The error-state dynamics function $f_{\mathrm{err},k}(\delta\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k)$, is straightforwardly defined as

$$
f_{\mathrm{err},k} \triangleq f_k \ominus f_{\mathrm{nom},k}
$$

where $\ominus$ denotes a generalized subtraction operator similar to $\oplus$ defined earlier.

The linearized covariance propagation step of the EKF requires computation of the following Jacobians.

$$F_k = \frac{\partial \boldsymbol{f}_{\text{err},k}(\delta\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k)}{\partial \delta\boldsymbol{x}_k}\Big|_{\substack{\delta\boldsymbol{x}_k=0 \\ \boldsymbol{w}_k=0}} \tag{4.7}$$

$$G_k = \frac{\partial \boldsymbol{f}_{\text{err},k}(\delta\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k)}{\partial \boldsymbol{w}_k}\Big|_{\substack{\delta\boldsymbol{x}_k=0 \\ \boldsymbol{w}_k=0}} \tag{4.8}$$

This involves calculus of rotations. The interested reader is referred to [74, 145] for further details. The nontrivial sub-blocks of F_k and G_k are documented in Appendix A.1.

4.5.4 Measurement Models & Calibration

This section details the measurement models for the various measurements applied to the error-state EKF, along with the calibration procedures necessary for the application of these measurements.

4.5.4.1 Inertial Measurements

IMUs measure the specific force and angular rate experienced by $\mathbf{b}$ relative to an inertial frame. If the centripetal force due to earth's rotation is absorbed in $\boldsymbol{g}^{\mathbf{n}}$, then the accelerometer and gyroscope measurements $\boldsymbol{z}^{\mathbf{b}}_{a,k}$ and $\boldsymbol{z}^{\mathbf{b}}_{\omega,k}$, respectively, are modeled as

$$\boldsymbol{z}^{\mathbf{b}}_{a,k} = R^{\mathbf{bn}}_k(\boldsymbol{a}^{\mathbf{n}}_k - \boldsymbol{g}^{\mathbf{n}}) + \boldsymbol{b}^{\mathbf{b}}_{a,k} + \boldsymbol{w}^{\mathbf{b}}_{a,k}$$

$$\boldsymbol{z}^{\mathbf{b}}_{\omega,k} = \boldsymbol{\omega}^{\mathbf{b}}_k + R^{\mathbf{bn}}_k\boldsymbol{\omega}^{\mathbf{n}}_{\mathbf{e}} + \boldsymbol{b}^{\mathbf{b}}_{\omega,k} + \boldsymbol{w}^{\mathbf{b}}_{\omega,k}$$

where $\boldsymbol{a}^{\mathbf{n}}_k$ is the true acceleration of the IMU in the $\mathbf{n}$ frame, which double-integrates to position deviation, and $\boldsymbol{\omega}^{\mathbf{b}}_k$ is the true angular rate of the IMU in

the $\mathbf{n}$ frame, which integrates to orientation deviation. For low-quality IMUs, accelerometer and gyroscope scale factors may also need to be modeled. For the MEMS IMU used in this work, it was observed that modeling the scale factors did not yield any performance benefit.

The stochastic models for IMU white noise and random walk process are derived from the IMU specifications. In addition to such intrinsic calibration, extrinsic calibration of the IMU with respect to $\mathbf{s}$ is necessary for the application of other measurements expressed in $\mathbf{s}$. The vector $\boldsymbol{p}^{\mathbf{s}}_{\mathbf{sb}}$ from $\mathbf{s}$ to $\mathbf{b}$ is taken to be known from the mechanical specification since this is not strongly observable from the available measurements. It is, however, important to estimate any deviations from the mechanically specified orientation $\bar{\boldsymbol{q}}^{\mathbf{sb}}$ between $\mathbf{b}$ and $\mathbf{s}$, since even sub-degree errors in the IMU orientation relative to $\mathbf{s}$ may lead to substantial errors when multiplied with the lever arm to another sensor.

The orientation deviation of $\bar{\boldsymbol{q}}^{\mathbf{sb}}$ from truth, denoted $\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{sb}}$, can be effectively estimated when CDGNSS measurements from multiple antennas are available to the EKF, as will be discussed in Sec. 4.5.4.2. Accordingly, the state vector $\delta\boldsymbol{x}_k$ is augmented with $\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{sb}}$ during clear-sky periods. It must be noted, however, that since the IMU is mounted near the line connecting the Sensorium's two GNSS antennas, only two of the three elements in $\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{sb}}$ are strongly observable. Any orientation deviation about the vector joining the two antennas is poorly unobservable, and must be constrained by construction. Also note that estimation of $\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{sb}}$ only need be performed once as long as all

sensors are rigidly mounted, and may not even be necessary if the mechanical tolerances are acceptably small.

4.5.4.2 CDGNSS Measurements

CDGNSS offers cm-accurate position measurements under all weather conditions, but typically offers reduced solution availability in deep urban environments. This chapter takes the approach of incorporating CDGNSS measurements in the localization engine whenever they are available, while being capable of maintaining the required lane-level accuracy over long CDGNSS outages in deep urban canyons. In essence, the approach developed in this chapter leverages CDGNSS for periodic or one-time intrinsic and extrinsic calibration of other on-board sensors, and relies on these sensors for accurate localization when CDGNSS is unavailable.

Signals captured from the two GNSS antennas on the Sensorium are processed together with those from a nearby reference station to provide nearly-independent three-dimensional position measurements of the antennas in the $\mathbf{n}$ frame. The position measurement for antenna $\mathbf{a}_i,\ i \in \{0, 1\}$ is modeled as

$$\boldsymbol{z}^{\mathbf{n}}_{\mathbf{a}_i,k} = \boldsymbol{p}^{\mathbf{n}}_k + R^{\mathbf{nb}}_k R^{\mathbf{bs}} \boldsymbol{p}^{\mathbf{s}}_{\mathbf{ba}_i} + \boldsymbol{e}_{\mathbf{a}_i,k} \tag{4.9}$$

where $\boldsymbol{e}_{\mathbf{a}_i,k}$ is the CDGNSS measurement noise. The vector $\boldsymbol{p}^{\mathbf{s}}_{\mathbf{ba}_i}$ from $\mathbf{b}$ to the antenna $\mathbf{a}_i$, expressed in $\mathbf{s}$, is available from the mechanical specification. As discussed above, $R^{\mathbf{bs}}$ may be taken to be the same as $\bar{R}^{\mathbf{bs}}$ from the mechanical specification, or may be further calibrated by augmenting the state with $\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{sb}}$.

Additionally, the error-state EKF requires the Jacobian of the measurement model with respect to the error state:

$$H_{\mathbf{a}_i,k} \triangleq \left.\frac{\partial \boldsymbol{z}^{\mathbf{n}}_{\mathbf{a}_i,k}}{\partial \delta \boldsymbol{x}_k}\right|_{\substack{\delta \boldsymbol{x}_k = 0 \\ \boldsymbol{e}_{\mathbf{a}_i,k}=0}} = \left.\frac{\partial \boldsymbol{z}^{\mathbf{n}}_{\mathbf{a}_i,k}}{\partial \boldsymbol{x}_k}\right|_{\substack{\boldsymbol{x}_k = \boldsymbol{x}_{\mathrm{nom},k} \\ \boldsymbol{e}_{\mathbf{a}_i,k}=0}} \cdot \left.\frac{\partial \boldsymbol{x}_k}{\partial \delta \boldsymbol{x}_k}\right|_{\substack{\delta \boldsymbol{x}_k = 0 \\ \boldsymbol{e}_{\mathbf{a}_i,k}=0}}$$

The nontrivial sub-blocks of $H_{\mathbf{a}_i,k}$ are documented in Appendix A.1.

4.5.4.3 Radar Range Rate & Bearing Measurements

The range rate and bearing measurements from automotive radars provide a valuable velocity constraint for inertial navigation. Importantly, the frequency modulated continuous wave (FMCW) signal used in automotive radars provides instantaneous range rate measurements to the detected targets, i.e., target tracking and/or matching across cluttered radar scans is not necessary to obtain and apply this measurement.

The relative velocity of a stationary target with respect to $\mathbf{r}_i$ is given by the negative of the velocity with respect to $\mathbf{n}$ of the ith radar, expressed in $\mathbf{r}_i$, written $-\boldsymbol{v}^{\mathbf{r}_i}_{\mathbf{r}_i,k}$, as shown in Fig. 4.5. Assuming that the radar only detects targets in the two-dimensional plane of the linear phased array, the range rate measurement is modeled as

$$\dot{r}_{ij,k} = \begin{bmatrix} \sin\theta_{ij,k} \\ -\cos\theta_{ij,k} \\ 0 \end{bmatrix}^{\top} R^{\mathbf{r}_i\mathbf{s}} R^{\mathbf{sb}} \left(R^{\mathbf{bn}}_k \boldsymbol{v}^{\mathbf{n}}_k + \left(\boldsymbol{\omega}^{\mathbf{b}}_k \times R^{\mathbf{bs}} \boldsymbol{p}^{\mathbf{s}}_{\mathbf{br}_i} \right) \right) \tag{4.10}$$

where the vector $\boldsymbol{p}^{\mathbf{s}}_{\mathbf{br}_i}$ and the radar orientation $R^{\mathbf{r}_i\mathbf{s}}$ may be taken from the mechanical specifications. Note that unlike typical measurement models where the right-hand side is composed of quantities that are either known or are

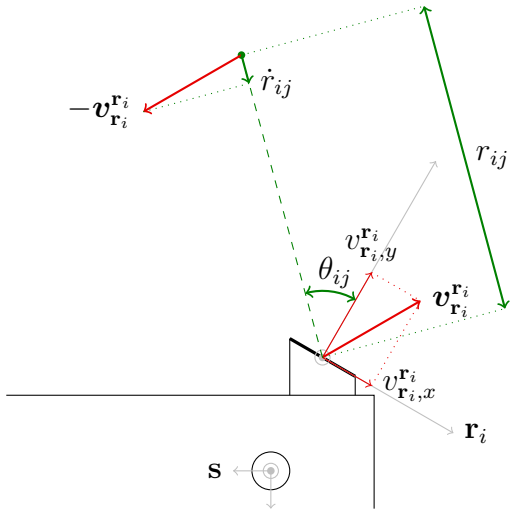


Figure 4.5: A visual description of the radar range rate measurement model. Quantities labeled in green are measured by the radar. The relative velocity of a stationary target with respect to $\mathbf{r}_i$ is the negative of the velocity with respect to $\mathbf{n}$ of the ith radar, expressed in $\mathbf{r}_i$, written $-\boldsymbol{v}_{\mathbf{r}_i,k}^{\mathbf{r}_i}$. The measured radial velocity $\dot{r}_{ij}$ of the jth stationary target is the projection of $-\boldsymbol{v}_{\mathbf{r}_i,k}^{\mathbf{r}_i}$ onto the line-of-sight direction between the ith radar and the jth target.

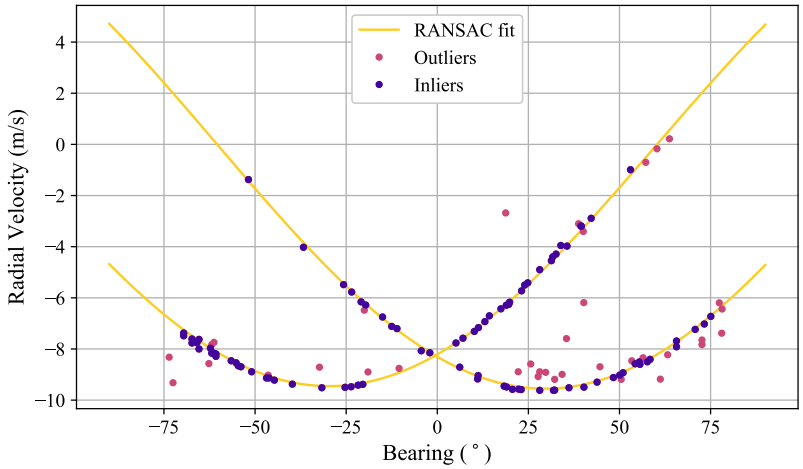


Figure 4.6: Example results of the RANSAC operation on radar range rate and bearing measurements. The two yellow sinusoidal curves represent the RANSAC-predicted radial velocities for the port and starboard radars from Fig. 4.4 as a function of the bearing. With a threshold of $0.2\,\mathrm{m\,s^{-1}}$, RANSAC considers violet dots as inliers and magenta dots as outliers. Note that the radial velocity magnitude is maximized at $-30°$ and $30°$ for the port and starboard radars, respectively, in agreement with the mounting angles of these radars on the vehicle.

being estimated, 4.10 has *measured* quantities $\theta_{ij,k}$ on the right-hand side of the equation. This implies that any errors in the bearing measurements will not be accounted for if the range rate measurements are modeled in the EKF as shown.

The application of range rate constraints comes with two major challenges. First, individual radar scans contain a number of spurious targets as discussed in Sec. 4.2. Second, automotive phased-array radars exhibit poor bearing resolution and accuracy, and this is further exacerbated by the unusual range rate measurement model described above. Both of these challenges are

addressed by pre-processing the range rate and bearing measurements with a RANSAC (random sample consensus) routine that estimates a best-fit two-dimensional radar velocity model to the radar measurements. In particular, with N detected targets, the RANSAC operation finds a robust solution to the following system of equations:

$$\begin{bmatrix} \dot{r}_{i0} \\ \vdots \\ \dot{r}_{iN} \end{bmatrix} = \begin{bmatrix} \sin\theta_{i0} & -\cos\theta_{i0} \\ \vdots & \vdots \\ \sin\theta_{iN} & -\cos\theta_{iN} \end{bmatrix} \begin{bmatrix} v^{\mathbf{r}_i}_{\mathbf{r}_i,x} \\ v^{\mathbf{r}_i}_{\mathbf{r}_i,y} \end{bmatrix} \tag{4.11}$$

while eliminating the $(\dot{r}_{ij}, \theta_{ij})$ pairs that may be outliers. Example results from the RANSAC procedure are shown in Fig. 4.6. Ultimately, the solution to 4.11 is applied as a measurement to the EKF with the following measurement model:

$$\begin{aligned} \boldsymbol{z}^{\mathbf{r}_i}_{\mathbf{r}_i,k} &\triangleq \begin{bmatrix} v^{\mathbf{r}_i}_{\mathbf{r}_i,x} \\ v^{\mathbf{r}_i}_{\mathbf{r}_i,y} \end{bmatrix}_k \\ &= \left[R^{\mathbf{r}_i\mathbf{s}} R^{\mathbf{sb}} \left(R^{\mathbf{bn}}_k \boldsymbol{v}^{\mathbf{n}}_k + \left(\boldsymbol{\omega}^{\mathbf{b}}_k \times R^{\mathbf{bs}} \boldsymbol{p}^{\mathbf{s}}_{\mathbf{br}_i} \right) \right) \right]_{[0,1]} + \boldsymbol{e}_{\mathbf{r}_i,k} \end{aligned}$$

where the subscript $[0, 1]$ denotes the first two elements of the three-element vector. Parts of the Jacobian of this measurement model with respect to the EKF error-state are documented in Appendix A.1.

4.5.4.4 Ground Vehicle Dynamics Constraints

Under nominal driving conditions, a ground vehicle respects dynamical constraints which can be leveraged as measurements to the EKF. This chapter incorporates near-zero sideslip and vertical velocity constraints, commonly referred to as nonholonomic constraints (NHC), as well as zero-speed updates (ZUPT). The measurement models for these constraints are described below.

Nonholonomic Constraints (NHC) The application of NHC is based on the following assumptions:

[T1] There exists a fixed center of rotation, taken to be the origin of $\mathbf{v}$, about which the vehicle rotates when a steering control input is applied.

[T2] When a zero steering input is applied, the vehicle only moves in the $\mathbf{v}_y$ direction. This holds by definition of $\mathbf{v}$.

[T3] The vehicle does not slip sideways or leave the surface of the road.

When the above assumptions hold, it follows that the velocity of the vehicle, when expressed in $\mathbf{v}$, is zero in the $\mathbf{v}_x$ and $\mathbf{v}_z$ directions at all times. In practice, however, these assumptions only hold approximately. Accordingly, the zero sideslip and vertical velocity constraints are applied as *soft* constraints in the form of measurements with an associated measurement error covariance. The NHC is modeled as

$$\begin{aligned} \mathbf{0}_{2\times 1} &\triangleq \boldsymbol{z}^{\mathbf{v}}_{\mathrm{nhc},k} && (4.12)\\ &= [\boldsymbol{v}^{\mathbf{v}}_k]_{[0,2]} + \boldsymbol{e}_{\mathrm{nhc},k} \\ &= \left[R^{\mathbf{vs}} R^{\mathbf{sb}} \left(R^{\mathbf{bn}}_k \boldsymbol{v}^{\mathbf{n}}_k + \left(\boldsymbol{\omega}^{\mathbf{b}}_k \times R^{\mathbf{bs}} \boldsymbol{p}^{\mathbf{s}}_{\mathbf{bv}}\right)\right)\right]_{[0,2]} + \boldsymbol{e}_{\mathrm{nhc},k} && (4.13) \end{aligned}$$

where $\boldsymbol{p}^{\mathbf{s}}_{\mathbf{bv}} = \boldsymbol{p}^{\mathbf{s}}_{\mathbf{bs}} + \boldsymbol{p}^{\mathbf{s}}_{\mathbf{sv}}$ and $R^{\mathbf{vs}}$ are parts of the extrinsic calibration between $\mathbf{v}$ and $\mathbf{s}$. Precise manual measurement of $\boldsymbol{p}^{\mathbf{s}}_{\mathbf{sv}}$ and $R^{\mathbf{vs}}$ is challenging. First, it is not obvious where the origin of $\mathbf{v}$ lies, though the center of line connecting the two rear axles might be a reasonable guess. Second, it would be challenging to measure, for example, the pitch of the Sensorium relative to the plane of

the vehicle chassis. Accordingly, this chapter takes a data-driven approach to extrinsic calibration of $\mathbf{v}$.

Once again, the extrinsic calibration technique relies on clear-sky periods with good CDGNSS availability, such that the nominal state estimates of $\boldsymbol{v}_k^{\mathbf{n}}$, $\boldsymbol{q}_k^{\mathbf{nb}}$, and $\boldsymbol{b}_{\omega,k}^{\mathbf{b}}$ are close to their true values. Furthermore, calibration begins with coarse initial guesses of $R^{\mathbf{vs}}$ and $\boldsymbol{p}_{\mathbf{sv}}^{\mathbf{s}}$, denoted $\bar{R}^{\mathbf{vs}}$ and $\bar{\boldsymbol{p}}_{\mathbf{sv}}^{\mathbf{s}}$, respectively, and attempts to estimate the orientation deviation $\boldsymbol{\eta}_{\mathbf{vs}}^{\mathbf{s}}$ and lever arm deviation $\delta\boldsymbol{p}_{\mathbf{sv}}^{\mathbf{s}}$ with respect to these. With other quantities assumed known, 4.13 may be rewritten as

$$\begin{aligned}\boldsymbol{e}_{\mathrm{nhc},k} &= \left[\left(\bar{R}^{\mathbf{vs}} \oplus \boldsymbol{\eta}_{\mathbf{vs}}^{\mathbf{s}}\right)\left(\boldsymbol{v}_k^{\mathbf{s}} + \left(\boldsymbol{\omega}_k^{\mathbf{s}} \times \left(\bar{\boldsymbol{p}}_{\mathbf{bv}}^{\mathbf{s}} + \delta\boldsymbol{p}_{\mathbf{bv}}^{\mathbf{s}}\right)\right)\right)\right]_{[0,2]} \\ &\triangleq \boldsymbol{h}_{\mathrm{nhc},k}(\boldsymbol{\eta}_{\mathbf{vs}}^{\mathbf{s}}, \delta\boldsymbol{p}_{\mathbf{bv}}^{\mathbf{s}})\end{aligned}$$

This model is nonlinear in $\boldsymbol{\eta}_{\mathbf{vs}}^{\mathbf{s}}$, and may be solved as a nonlinear least squares problem, e.g., with the Gauss-Newton method. The Jacobian of $\boldsymbol{h}_{\mathrm{nhc},k}$ evaluated at $\boldsymbol{\eta}_{\mathbf{vs}}^{\mathbf{s}} = \mathbf{0}$ and $\delta\boldsymbol{p}_{\mathbf{bv}}^{\mathbf{s}} = \mathbf{0}$ is composed of

$$\begin{aligned}\frac{\partial \boldsymbol{h}_{\mathrm{nhc},k}}{\partial \boldsymbol{\eta}_{\mathbf{vs}}^{\mathbf{s}}} &= \left[\left(\boldsymbol{v}_k^{\mathbf{s}} + \boldsymbol{\omega}_k^{\mathbf{s}} \times \bar{\boldsymbol{p}}_{\mathbf{bv}}^{\mathbf{s}}\right)^\top \otimes \left[\bar{R}^{\mathbf{vs}}\right]_{[(0,2),(:)]}\right] \begin{bmatrix}[-\hat{\mathbf{i}}]_\times \\ [-\hat{\mathbf{j}}]_\times \\ [-\hat{\mathbf{k}}]_\times\end{bmatrix} \\ \frac{\partial \boldsymbol{h}_{\mathrm{nhc},k}}{\partial \delta\boldsymbol{p}_{\mathbf{bv}}^{\mathbf{s}}} &= \left[\bar{R}^{\mathbf{vs}}\right]_{[(0,2),(:)]} [\boldsymbol{\omega}_k^{\mathbf{s}}]_\times\end{aligned}$$

where $\otimes$ denotes the Kronecker product, subscript $[(0,2),(:)]$ denotes selection of the first and third rows of a matrix, $[\cdot]_\times$ denotes the skew-symmetric cross-product matrix corresponding to the 3-element argument, and $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ denote the cardinal unit vectors. To make the system observable, measurements

from multiple epochs must be stacked and solved as a batch. Additionally, the nonlinear problem must be iteratively linearized and solved until convergence.

Zero-Speed Update (ZUPT) The ZUPT constraint is another valuable measurement that limits odometric drift, especially in situations where the platform makes frequent stops. The measurement model for ZUPT is trivially written as

$$\begin{aligned} \mathbf{0}_{3\times 1} &\triangleq \boldsymbol{z}^{\mathbf{v}}_{\text{zupt},k} \\ &= R^{\mathbf{vs}} R^{\mathbf{sb}} R^{\mathbf{bn}}_{k} \boldsymbol{v}^{\mathbf{n}}_{k} + \boldsymbol{e}_{\text{zupt},k} \end{aligned} \tag{4.14}$$

The primary challenge of applying ZUPT is detection of epochs where this constraint is valid. Importantly, this condition must be detected independently from the EKF state estimate, e.g., by inspection of the raw IMU measurements. In theory, it is not possible to make any claims about zero speed based on acceleration and/or angular rate data, since IMU measurements of a vehicle moving with a constant velocity and orientation must be indistinguishable from those of a stationary vehicle. In practice, however, the IMU measurements exhibit a distinct behavior when the vehicle is in motion, e.g., due to road roughness and vehicle vibrations. Prior work has made use of these *artifacts* to detect stationary periods. This chapter follows the angular rate energy method from [144] for ZUPT detection. In practice, if wheel odometry data are available from the vehicle CAN bus, as is common in most modern vehicles, then ZUPT detection can be performed trivially and with high reliability.

An observant reader might wonder why ZUPT is not applied directly to $\boldsymbol{v}_k^{\mathbf{n}}$ in 4.14. The advantage of applying ZUPT in $\mathbf{v}$ is that a tighter zero-speed constraint can be reliably applied in the lateral and vertical directions.

4.5.5 Batch Smoothing & Update

Real-time estimates of the vehicle pose trajectory obtained from the EKF may be used to string together individual scans and perform a radar batch measurement update. However, since these data are processed batches, it is desirable to perform backward smoothing over the short duration of the batch. Backward smoothing enforces the dynamics function backwards in time, ironing out any large jumps that may have occurred in the EKF forward pass.

Accordingly, the batch smoother component in Fig. 4.3 stacks all inertial measurements and snapshots of the estimator state over the duration of the batch. When the batch is ready to be processed for correlation, backward smoothing is enforced with the inertial measurements as control inputs. The smoothing formulation in this case is somewhat more complicated than usual [132] due to nonlinear backward dynamics and the error-state formulation. Details on nonlinear error-state Rauch-Tung-Striebel smoothing are provided in Appendix A.2.

The correlation peak search region is taken to be $\pm 5\,\mathrm{m}$ and $\pm 3^{\circ}$. The 3-DoF pose offset $\widehat{\boldsymbol{\Theta}}$ from radar batch correlation is applied as horizontal position and heading measurements to the EKF. Outliers from batch correlation are

excluded in the EKF based on a χ^2-test on the normalized innovation squared (NIS) [11].

4.6 Experimental Results

The radar-inertial positioning system of Fig. 4.3 was evaluated experimentally using the dataset described in [111], collected during approximately 1.5 h of driving on two separate days in and around the urban center of Austin, TX. This section presents the evaluation results.

4.6.1 Dataset

Fig. 4.7 shows the route followed by the sensor-instrumented vehicle on Thursday, May 9, 2019 (in blue) and Sunday, May 12, 2019 (in red). The test route combs through every street in the Austin, TX downtown area, since such environments are the most challenging for CDGNSS-based positioning [67] and would benefit the most from multi-sensor all-weather positioning. The route was driven once on a weekday and again on the weekend to evaluate robustness of the proposed map-based approach to changes in the traffic and parking patterns. Note that the final part of the route (the north-east segment) was different on the two days, preventing the use of a map-based positioning approach. This section of the test route has been omitted from the evaluation results.

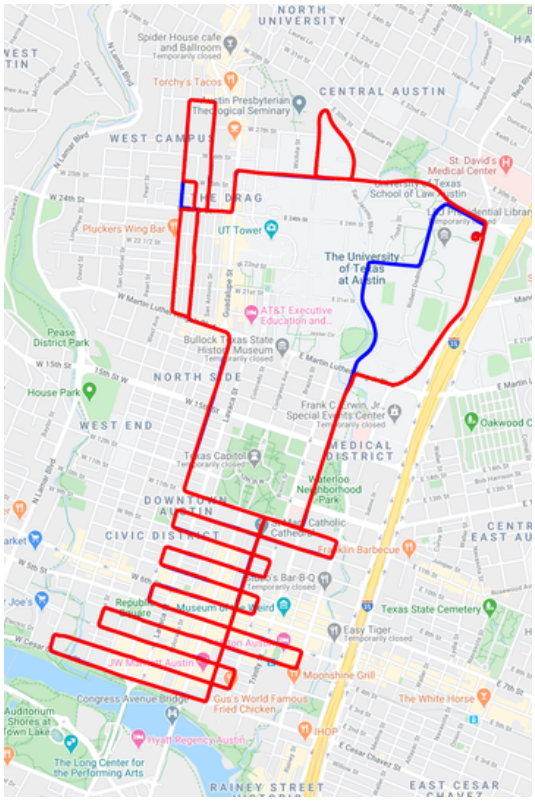


Figure 4.7: Test route through The University of Texas west campus and Austin downtown. These areas are the most challenging for precise GNSS-based positioning and thus would benefit the most from radar-based positioning. The route was driven once on a weekday and again on the weekend to evaluate robustness of the radar map to changes in traffic and parking patterns. Red is the mapping run (May 12), blue is the localization run (May 9). A prior map is not available in the visible blue areas.

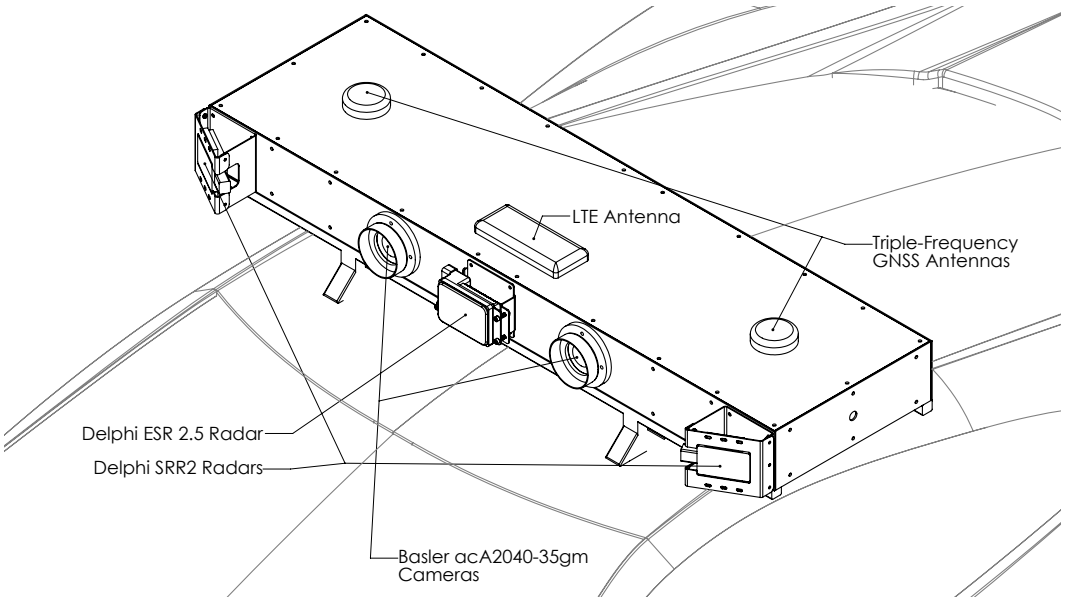


Figure 4.8: The University of Texas Sensorium is an integrated platform for automated and connected vehicle perception research. It includes three automotive radar units, one electronically-scanning radar (ESR) and two short-range radars (SRR2s); stereo visible light cameras; automotive- and industrial-grade IMUs; a dual-antenna, multi-frequency software-defined GNSS receiver; 4G cellular connectivity; and a powerful internal computer. An iXblue ATLANS-C CDGNSS-disciplined INS (not shown) is mounted at the rear of the platform to provide the ground truth trajectory.

4.6.1.1 Sensors

Designed for connected and automated vehicle research, the Sensorium, shown in Fig. 4.8 is a self-contained sensor housing that can be mounted atop any standard passenger vehicle. Two Antcom G8Ant-3A4TNB1 triple-frequency patch antennas are flush-mounted in the cross-track direction on the Sensorium's upper plate, separated by just over one meter. The antennas' signals are routed to a unified radio frequency (RF) front end whose output intermediate frequency (IF) samples are processed in real-time (to within less than 10 ms latency) by the Sensorium's onboard computer. The samples are also stored to disk for post-processing.

The Sensorium features a front-facing stereo camera rig composed of two Basler acA2040-35gm cameras that capture synchronous stereo image pairs when triggered by a signal tied to the GNSS front-end's sampling clock. The images are captured in grayscale at 10 frames per second and timestamped by the Sensorium's computer. The cameras are configured to automatically adjust the exposure time based on lighting, while the focal length, focus, and aperture are held fixed, having been adjusted physically prior to capture.

The Sensorium is also equipped with two types of automotive radars: one Delphi electronically-scanning radar (ESR) in the middle and two Delphi short-range radars (SRR2s) on the two sides. Both the ESR and the SRR2 are commercially available; similar radars are available on economy-class consumer vehicles. Fig. 4.9 shows the coverage patterns for the three radar units. The ESR provides simultaneous sensing in a narrow ($\pm 10^\circ$) long-range (175 m)

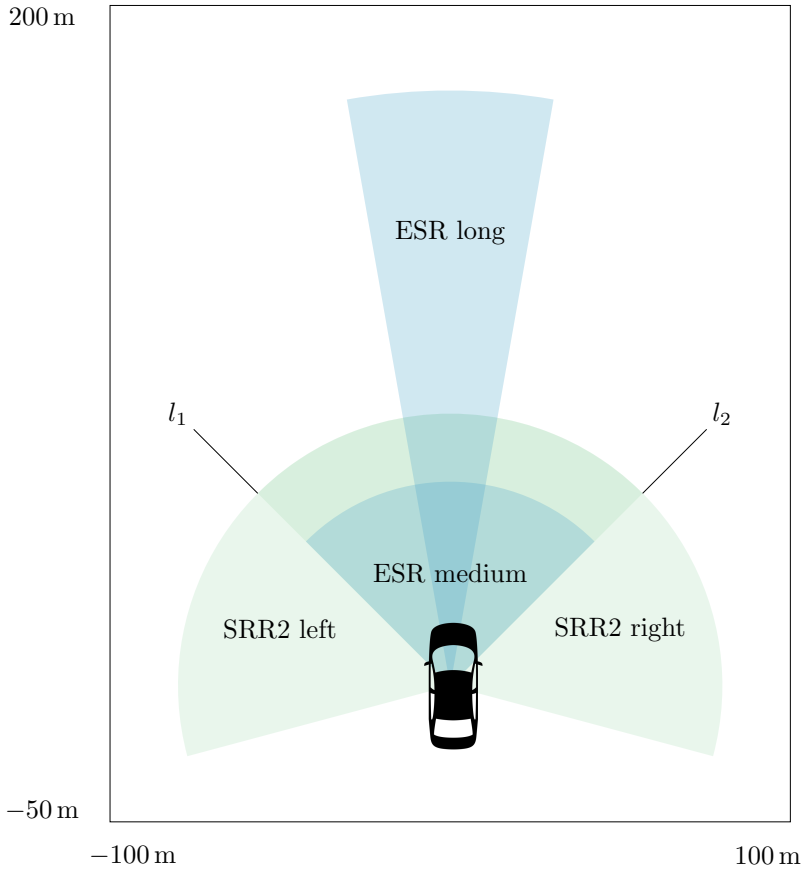


Figure 4.9: Coverage patterns for the three Sensorium radar units. The ESR provides simultaneous sensing in a narrow (±10°) long-range (175 m) coverage area and a wider (±45°) medium-range (60 m) area. The SRR2 units each have a coverage area of ±75° and 80 m. The line l_1 marks the left-most extent of the right SRR2's field of view. Similarly, l_2 marks the right-most extent of the left SRR2's field of view. Each SRR2 is installed facing outward from the centerline at an angle of 30°.

coverage area and a wider (±45°) medium-range (60 m) area. The SRR2 units each have a coverage area of ±75° and 80 m. Each SRR2 is installed facing outward from the center-line at an angle of 30°. The Sensorium's onboard computer timestamps and logs the radar returns from the three radar units.

The LORD MicroStrain 3DM-GX5-25 MEMS IMU is an industrial-grade inertial sensor that acts as the core sensor of the localization pipeline. The IMU provides temperature-compensated accelerometer and gyroscope readings at 100 Hz.

4.6.1.2 Ground-Truth Trajectory

The ground-truth position and orientation trajectory for the data are generated with the iXblue ATLANS-C, a high-performance CDGNSS coupled fiber-optic gyroscope INS. The post-processed position solution obtained from the ATLANS-C is decimeter-accurate throughout the dataset.

4.6.1.3 Dataset Splits

With a limited amount of field data available for development and evaluation, it is critical to ensure that the proposed positioning technique does not overfit this particular dataset. Accordingly, the data used in the development of the algorithms were restricted to a fixed 30 min segment, where the prior radar map was constructed with radar measurements from May 9 and localization was performed with radar, inertial, and CDGNSS measurements from May 12. In contrast, during evaluation the full 62 min of data were used,

and the mapping and localization datasets were inverted, i.e., the prior map was constructed with radar measurements from May 12, and localization was performed with all sensor data from May 9. The algorithms have not been modified to maximize the performance over the evaluation dataset.

4.6.2 Prior Radar Mapping

The first step to radar-map-based localization is the generation of a radar map point cloud. Radar scans collected from the May 12, 2019 drive were aggregated to create a map with the benefit of the ATLANS-C ground-truth trajectory. In a practical system, the radar map may be generated during favorable conditions for optical sensors such as cameras and lidar, such that the mapping vehicle can accurately track its pose. Additionally, the mapping process may be crowed-sourced from consumer vehicles [112, 113]. The map point cloud is stored in a k-d tree for efficient querying during localization.

Two implementation notes are in order here. First, automotive radar clutter is especially intense when the vehicle is stationary. Accordingly, radar range measurements obtained when the vehicle was moving slower than $1\,\mathrm{m\,s^{-1}}$ were discarded for both mapping and localization. This implies that radar correlation measurements were only available during periods when the vehicle was moving faster than $1\,\mathrm{m\,s^{-1}}$. Second, it was observed that radar returns far from the vehicle are mostly clutter and have negligible resemblance to the surrounding structure. Radar returns with range larger than $50\,\mathrm{m}$ were discarded for both the map and batch PHDs. It is noted that these two

parameters have not been optimized to produce the smallest estimation errors; instead they have been fixed based on visual inspection.

4.6.3 Offline Calibration

Extrinsic calibration among the IMU frame $\mathbf{b}$, the Sensorium frame $\mathbf{s}$, and the vehicle frame $\mathbf{v}$ was performed offline with 125 s of sensor data with CDGNSS availability. While it is possible to estimate the calibration parameters online, it may not be desirable to do so if these parameters are not expected to change over time.

The orientation deviation $\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{sb}}$ between the IMU body frame and the Sensorium frame was calibrated for the localization dataset, as described in Sec. 4.5.4.1. With two GNSS antennas, only two out of the three DoFs in $\boldsymbol{\eta}^{\mathbf{s}}_{\mathbf{sb}}$ are observable. Accordingly, the orientation deviation around $\mathbf{b}_x$, which is mostly unobservable, was tightly constrained to the initial guess of zero. The deviations around $\mathbf{b}_y$ and $\mathbf{b}_z$ rapidly converged to sub-degree offsets from the mechanical specification.

Extrinsic calibration between $\mathbf{v}$ and $\mathbf{s}$ was similarly estimated over the 125 s period as detailed in Sec. 4.5.4.4.

The commercial automotive radars on the Sensorium do not offer any mechanism to synchronize their scans with an external reference clock. Analysis of the radar range rate residuals in the EKF showed clear evidence of latency between the data logging timestamp and the true scan times. Accordingly, radar latency calibration was performed offline with a best fit approach.

4.6.4 Implementation Notes

A few implementation- and dataset-specific notes relating to the localization pipeline are documented below.

CDGNSS Measurements & Outages The CDGNSS position measurements used in this evaluation are in fact the output of the post-processed ground-truth system, i.e., these measurements have not been obtained from an unaided CDGNSS receiver. While this is not ideal for realistic evaluation, the evaluation results presented herein do not mislead because, first, CDGNSS measurements are only applied for a 125 s period for initial calibration, and second, any commercial CDGNSS receiver would be able to generate similar cm-accurate position solutions in the clear-sky region where the CDGNSS measurements were applied.

Measurement Noise Correlation Observations from the field data revealed that the measurement noise in the radar range rate measurements is not independent between consecutive radar scans. This is problematic since the EKF applied assumes each measurement to have errors that are uncorrelated in time. Accordingly, the radar range rate measurements were decimated to 1 Hz such that the measurements were spaced out by roughly the decorrelation time of the measurement noise. A more principled approach to this problem is to augment the state vector with states to pre-whiten the measurements. But this approach was empirically observed to not outperform the straightforward

Table 4.1: Tuning parameters involved in the radar-inertial localization pipeline

Minimum speed for valid radar range	$1\,\mathrm{m\,s^{-1}}$
Maximum valid radar range	$50\,\mathrm{m}$
Minimum RANSAC inliers	10
Minimum fraction of RANSAC inliers	0.65
$v^{\mathbf{r}_i}_{\mathbf{r}_i,x}$ (broadside) standard deviation	$0.2\,\mathrm{m\,s^{-1}}$
$v^{\mathbf{r}_i}_{\mathbf{r}_i,y}$ (boresight) standard deviation	$0.1\,\mathrm{m\,s^{-1}}$
$v^{\mathbf{v}}_{\mathrm{nhc},x}$ (lateral) standard deviation	$0.1\,\mathrm{m\,s^{-1}}$
$v^{\mathbf{v}}_{\mathrm{nhc},z}$ (vertical) standard deviation	$0.2\,\mathrm{m\,s^{-1}}$

measurement decimation, while introducing additional complexity and tuning parameters.

Similarly, the NHC and ZUPT measurements can in theory be applied at every applicable IMU epoch. But to prevent correlated errors in these constraints (e.g., due to sideslip experienced while cornering) from making the EKF inconsistent, they are only applied at 1 Hz.

Filter Tuning Parameters The process noise covariance used in the EKF is derived from the IMU datasheet parameters [88, 168]. The measurement noise covariance associated with CDGNSS measurements is available directly from the ATLANS-C receiver. A few other measurement noise standard deviations and tuning parameters are documented in Table 4.1.

4.6.5 Localization Results with Perfect Odometry

This section evaluates the localization performance of the proposed method on the May 12, 2019 radar data for the (hypothetical) case in which odometric drift over the radar batch-of-scans interval is negligible. With decreasing quality of the odometry sensor(s), this assumption holds only over ever shorter batch intervals. Therefore, the performance of the algorithm is evaluated for a range of batch lengths.

4.6.5.1 Test procedure

A drift-free vehicle trajectory over a batch-of-scans interval is generated with the reference solution from the iXblue ATLANS-C. This trajectory is then artificially offset by a two-dimensional rigid transformation error. The translational error is distributed such that $\Delta \boldsymbol{t} \sim \mathcal{N}(\mathbf{0}, \sigma_t^2 I)$ with $\sigma_t = 2\,\mathrm{m}$, and the rotational error is distributed such that $\Delta\phi \sim \mathcal{N}(0, \sigma_\phi^2)$ with $\sigma_\phi = 3^\circ$. The proposed localization technique takes the erroneously-offset position and heading trajectory as the initial guess of the vehicle state.

The prior radar map point cloud in the vicinity of the initial guess of the vehicle position is retrieved with a query to the k-d tree. Additionally, the batch of body-frame radar returns is transformed to a common reference frame based on the erroneous trajectory. The goal is to align the two point clouds and thereby estimate the artificially-induced translational and rotational offset.

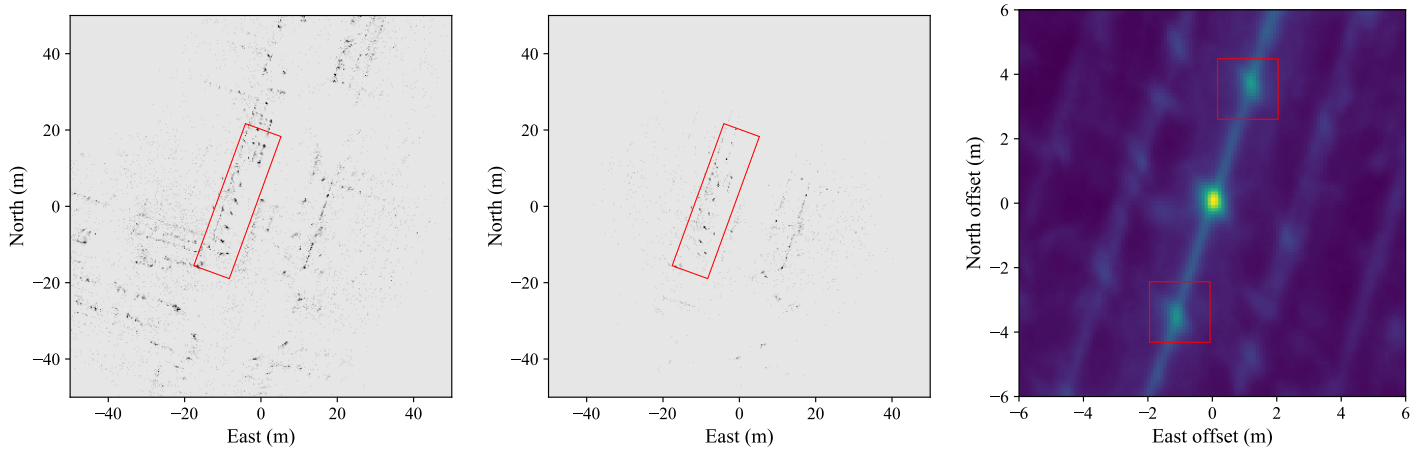
As a first step, the map and batch occupancy grids are generated based on the aggregated point clouds, following the procedure described in Sec. 4.4.

The extent of the occupancy grids is determined by the bounds of the area scanned by the radars during localization. Given the maximum range of radar returns considered here, the correlation region is typically restricted to $\pm 50\,\mathrm{m}$ around the provided position at the end of the batch. With a grid cell size of $10\,\mathrm{cm}$, the occupancy grid size in discrete coordinates is typically on the order of $n = 1000$. The translation search space is limited to $\pm 3\sigma_t = \pm\, 6\,\mathrm{m}$, resulting in $n_l = 120$. Similarly, the rotation search space is limited to $\pm 3\sigma_\phi = \pm\ 9^\circ$ with 1° steps, resulting in $m = 18$.

4.6.5.2 Drift-Free $5\,\mathrm{s}$ Batches

This section evaluates and analyzes the proposed method for a fixed batch length of $5\,\mathrm{s}$. Before diving into the quantitative analysis, it is interesting to inspect the example of a radar batch update shown in Fig. 4.10. For ease of visualization, the batch point cloud to be localized has already been adjusted for any translational or rotational offset from the ground truth. The occupancy grid estimated from the $5\,\mathrm{s}$ batch of scans is shown in Fig. 4.10b. Similarly, Fig. 4.10a shows the occupancy grid estimated from the map point cloud retrieved from the map database. Fig. 4.10c shows the cross-correlation between the batch and map occupancy grids. Given that the batch is already aligned with ground truth, one should expect the correlation peak to appear at $(0, 0)$ in Fig. 4.10c. The offset of the peak from $(0, 0)$ in this case would be the translational estimate error.

Two interesting features of the cross-correlation in Fig. 4.10c are worth



(a) (b) (c)

Figure 4.10: This figure shows an interesting example of radar-based urban positioning with the proposed method. Panel (a) shows the occupancy grid estimated from the prior map point cloud. Panel (b) shows the same for a 5 s batch of scans collected in the same region. For ease of visualization, the batch occupancy grid has already been aligned with the map occupancy grid. Panel (c) shows the cross-correlation between the batch and map occupancy grids at $\Delta\phi = 0°$. Given that no rotational or translational offset error has been applied to the batch, the correlation peak should appear at $(0, 0)$. The offset of the peak in panel (c) from $(0, 0)$ is the translational estimate error of the proposed method. Also note the increased positioning uncertainty in the along-track direction, and the two local correlation peaks (marked with red squares in panel (c)) due to the repeating periodic pattern of radar reflectors in the map and the batch (marked with red rectangles in panels (a) and (b)).

noting. First, the correlation peak decays slower in the along-track direction—in this case approximately aligned with the south-southwest direction. This is a general feature observed throughout the dataset, since most of the radar reflectors are aligned along the sides of the streets. Second, there emerge two local correlation peaks offset by ≈4 m along the direction of travel. These local peaks are due to the repeating periodic structure of radar reflectors in both the map and the batch occupancy grids. In other words, shifting the batch occupancy grid forward or backward along the vehicle trajectory by ≈4 m aligns the periodically-repeating reflectors in an off-by-one manner, leading to another plausible solution. Importantly, the uncertainty envelope of the initial position estimate can span several meters, encompassing both the global optimum and one or more local optima. This explains why gradient-based methods, which seek the nearest optimum, are poorly suited for use in the urban automotive radar environment.

The complementary cumulative distribution functions (CCDF), i.e., the fraction of epochs exceeding any given level of error, of the horizontal position and heading error magnitudes are plotted on a log scale in Fig. 4.11 for drift-free (hypothetical) 5 s batches. In 95% of epochs, the horizontal position error magnitude was no greater than 44 cm, and the heading error magnitude was no greater than 0.59°.

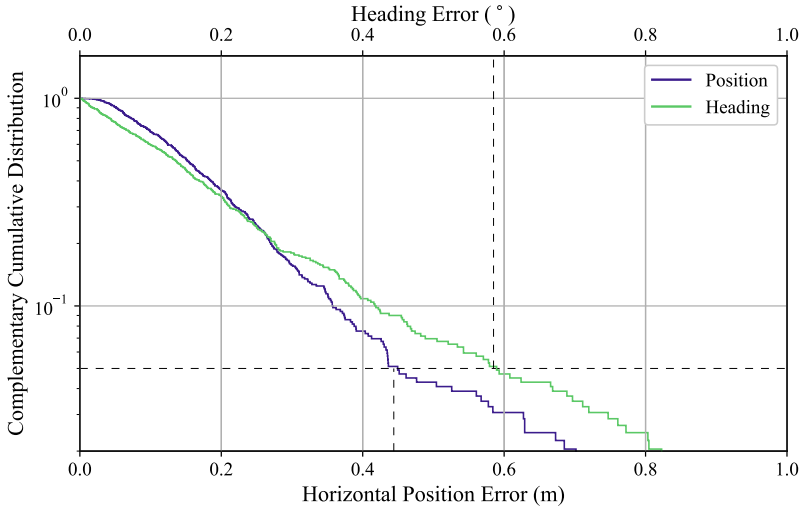


Figure 4.11: The complementary cumulative distribution (also known as a survival function) indicates how often (that is, in what fraction of 5 s epochs) the localization procedure in the text was found to exceed a given level of error. The logarithmic vertical scale makes the tails of the distribution, corresponding to outliers that may cause tracking errors, more visible. For drift-free (hypothetical) 5 s batches, the 95-percentile horizontal positioning error is observed to be 44 cm and the 95-percentile heading error is observed to be 0.59°.

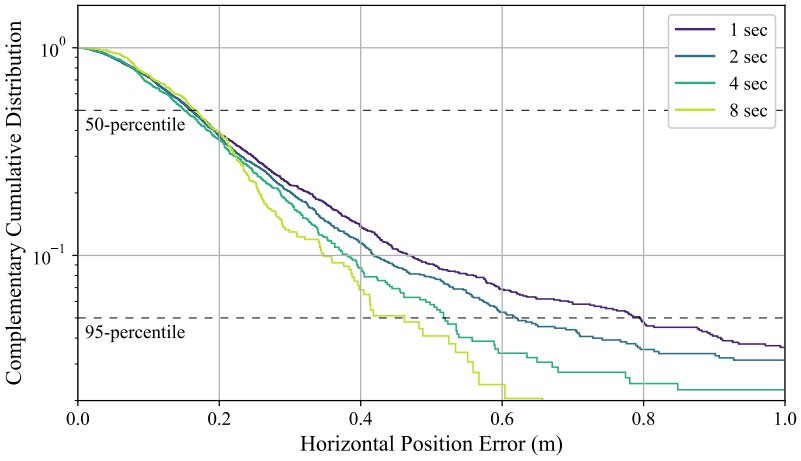


Figure 4.12: CCDFs for different drift-free batch lengths between 1 s and 8 s. The 50-percentile errors are similar for shorter and longer batch lengths, but the difference becomes more noticeable at higher percentiles.

4.6.5.3 Sensitivity to Batch Length

The assumption of negligible odometric drift over the batch-of-scans interval does not hold over 5 s for low-cost odometry sensors, but may hold for longer than 5 s for high-performance sensors. Thus, is important to evaluate the proposed localization technique's sensitivity to batch length.

Fig. 4.12 shows the CCDF for different drift-free batch lengths between 1 s and 8 s. As expected, the errors are smaller for longer batch lengths. It is interesting to note that the 50-percentile errors are similar for different batch lengths, but difference between the CCDFs becomes more pronounced at higher percentiles. This indicates that the shorter batch lengths are adequate in most cases, but the longer batches help contain the estimation error in a few cases. Such behavior must be taken into account in the integrity analysis of

the system. For example, while all batch lengths exhibit similar 50-percentile error behavior, batches shorter than 4 s cannot be used in applications with a 50 cm alert limit and integrity risk smaller than 0.05 per batch. On the other hand, in applications where 50-percentile error is the performance metric, it may be desirable to choose batches shorter than 4 s to relax the requirements on short term odometric performance.

4.6.6 Localization Results with Odometric Sensors

This section removes the assumption of perfect drift-free odometry and presents empirical error statistics obtained from field evaluation of the system in 4.3. The test scenario evaluated in this section is an extreme one: the vehicle starts off in a clear-sky environment with 125 s of CDGNSS availability, and subsequently all CDGNSS measurements are cut off for the next 3600 s of driving, during which the system must rely on radar and inertial sensing along with vehicle dynamical constraints to maintain an accurate estimate of its pose.

4.6.6.1 Performance with 4 s Radar Batches

Fig. 4.13 shows the east and north position error time histories from the test scenario described above. For the results presented in Fig. 4.13 and 4.14, a 4 s radar batch duration is chosen. In the first 125 s of clear-sky conditions with CDGNSS availability, the east and north position errors with respect to the ground truth are sub-decimeter, as expected. Over the subsequent

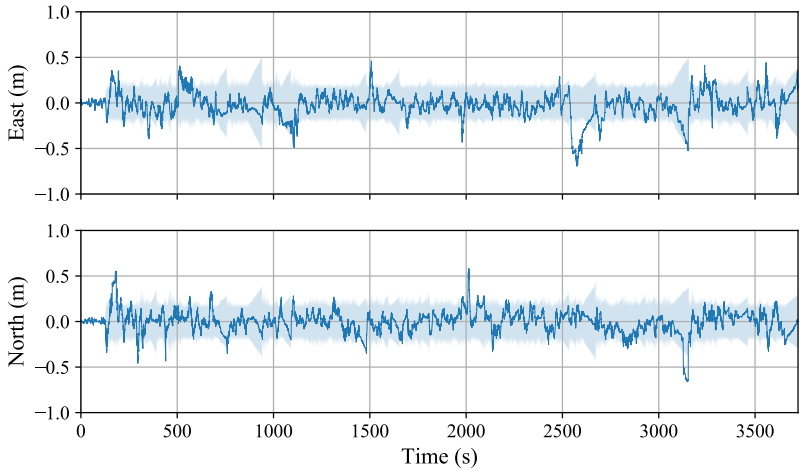


Figure 4.13: East and north position error time histories from field evaluation. In the first 125 s of clear-sky conditions with CDGNSS availability, the east and north position errors with respect to the ground truth are sub-decimeter, as expected. Over the subsequent 60 min of driving in and around the urban center of the city, the proposed method maintains sub-35-cm (95%) horizontal position errors. The horizontal position estimation errors are consistent with the predicted standard deviation from the EKF.

60 min of driving in and around the urban center of the city, the proposed method maintains sub-35-cm horizontal position errors (95%). The horizontal position estimation errors are consistent with the predicted standard deviation from the EKF. This is a remarkable result which shows that, given a prior radar map, lane-level-accurate horizontal positioning is achievable under zero-visibility GNSS-denied conditions with the types of sensors that are already available on commercial vehicles. Vertical position errors are not shown in Fig. 4.13 since these are not constrained by the two-dimensional radar batch correlation update. For ground vehicle applications, a digital elevation map can effectively constrain errors in altitude, if necessary.

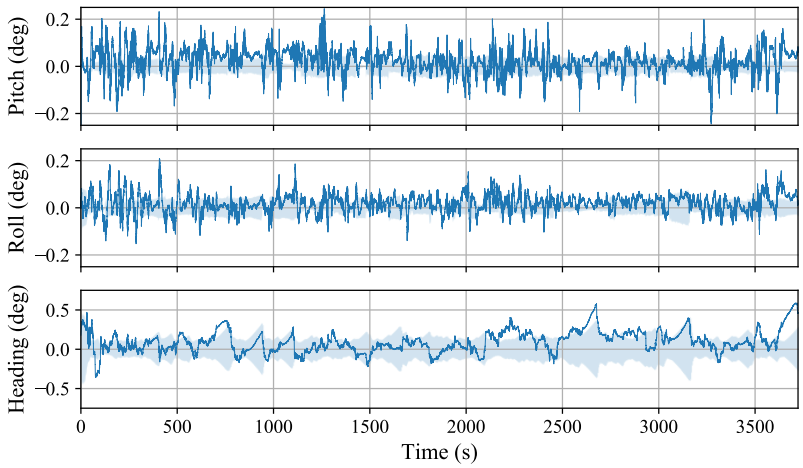


Figure 4.14: Vehicle orientation estimation errors from field evaluation. The proposed technique maintains vehicle heading estimates to within 0.5° of the ground truth throughout most of the dataset, and the errors are consistent with the predicted uncertainty. Roll and pitch estimation errors are smaller and stay within 0.2° of the ground truth.

Vehicle orientation estimation errors for the same scenario are shown in Fig. 4.14. Heading estimation error, shown in the bottom panel, is most important for ground vehicle applications. The proposed technique maintains vehicle heading estimates to within 0.5° of the ground truth throughout most of the dataset, and the errors are consistent with the predicted uncertainty. Heading error appears to be biased for a 600 s period beginning at 2500 s, but it must be noted that this is only a single realization of a random process with long time constants due to infrequent attitude changes. Roll and pitch estimation errors are smaller and stay within 0.2° of the ground truth. Better estimation of roll and pitch is expected since these are directly observable with the accelerometer measurements. The same phenomenon explains the substan-

tially shorter decorrelation times for roll and pitch errors as compared to the heading error. Finally, it is noted that the EKF is mildly inconsistent in regards to roll and pitch estimation errors. This suggests that the accelerometer white noise and bias stability characteristics claimed in the IMU datasheet [88] may be optimistic in field application.

4.6.6.2 Choosing a Radar Batch Length

The problem of choosing the duration of a radar batch during localization presents an interesting trade-off. On the one hand, longer batch durations are preferable because, as shown in Fig. 4.12, cross-correlation using a larger *patch* of the radar environment is more likely to produce a strong and unambiguous correlation peak. As mentioned earlier, the 50^{th} percentile horizontal position errors are similar for different batch lengths. The difference between the CCDFs becomes more pronounced at higher percentiles, implying that errors for shorter batch lengths have heavy tails. Recall that in the overall localization pipeline of Fig. 4.3, these errors will act as measurement errors in $\widehat{\boldsymbol{\Theta}}$. An EKF models measurement errors to be Gaussian, which is not a good model for heavy-tailed distributions. Accordingly, longer batch durations would appear preferable.

On the other hand, longer batches have several disadvantages. First, longer durations between batch measurement updates leads to larger odometric drift *between* updates, as well as poorer reconstruction of the radar batch itself. Second, some of the worst outliers due to shorter batch lengths may

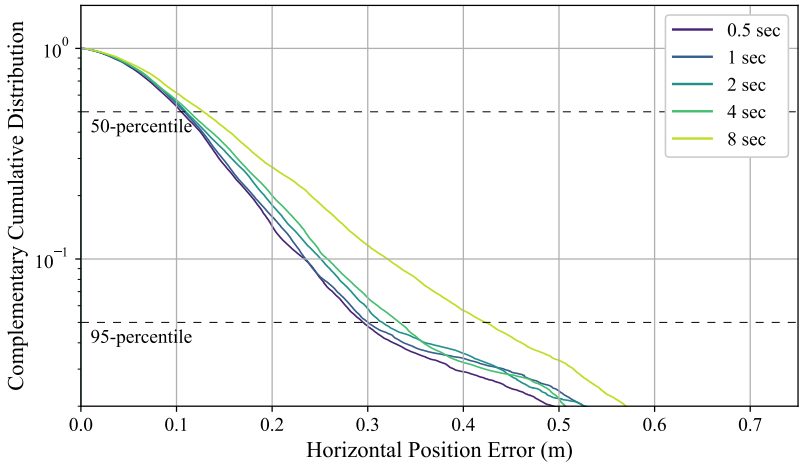


Figure 4.15: End-to-end effect of different batch lengths on horizontal positioning performance. Other than the longest batch length of 8 s, most batch lengths appear to perform similarly well, with 95$^{\text{th}}$-percentile horizontal position errors near 30 cm.

be rejected in the EKF based on the χ^2 NIS test, thus blunting the relative advantage of longer batches. Shorter batch lengths allow for a larger number of measurement updates to be performed per unit time, even if a few of those measurements may have to be rejected as outliers.

Fig. 4.15 reveals the end-to-end effect of different batch lengths. For a given batch length, its measurement error standard deviation was obtained from the corresponding CCDF in Fig. 4.12, i.e., the $\widehat{\boldsymbol{\Theta}}$ measurement standard deviation is smaller for longer batches. Interestingly, other than the longest batch length of 8 s, most batch lengths appear to perform similarly well, with 95-percentile horizontal position errors near 30 cm. Given the heavy-tailed nature of measurement noise distributions when working with very short batches (from Fig. 4.12), batch lengths from 2 to 4 s may be taken to be a good com-

promise.

4.7 Conclusion

A robust pipeline for all-weather sub-50-cm urban ground vehicle positioning has been proposed and evaluated. The positioning engine is based on commercially-available low-cost automotive radars, MEMS IMU, ground vehicle dynamics constraints, and, when available, precise GNSS measurements. Remarkably, it has been shown that given a prior radar map, lane-level-accurate horizontal positioning is achievable under zero-visibility GNSS-denied conditions with the types of sensors that are already available on commercial vehicles. In comparison with a post-processed ground truth trajectory, it was shown that during 60 min of GNSS-denied driving in the urban center of Austin, TX, the proposed pipeline has 95^{th}-percentile errors of 35 cm in horizontal position and 0.5° in heading. This is a significant development in the field of AGV localization, which has traditionally been based on sensors such as lidar and cameras that perform poorly in bad weather conditions.

Chapter 5

Accuracy Limits for Collaborative Globally-Referenced Digital Mapping with Standard GNSS

5.1 Abstract

Exchange of location and sensor data among connected and automated vehicles will demand accurate global referencing of the digital maps currently being developed to aid positioning for automated driving. This chapter explores the limit of such maps' globally-referenced position accuracy when the mapping agents are equipped with low-cost GNSS receivers performing standard code-phase-based navigation. The key accuracy-limiting factor is shown to be the asymptotic average of the error sources that impair standard GNSS positioning. Asymptotic statistics of each GNSS error source are analyzed through both simulation and empirical data to show that sub-50-cm accurate digital mapping is feasible in the horizontal plane after multiple mapping ses-

This chapter is based on: Lakshay Narula, J. Michael Wooten, Matthew J. Murrian, Daniel M. LaChapelle, and Todd E. Humphreys. Accurate collaborative globally-referenced digital mapping with standard GNSS. *Sensors*, 18(8), 2018.

sions with standard GNSS, but larger biases persist in the vertical direction.

5.2 Introduction

Localization is one of the primary operations that connected and automated vehicles must perform, both to navigate from one location to another and to interact with each other and with their surroundings within a mapped environment. Satellite-based navigation sensors have historically been the unrivalled sensor of choice for navigating from source to destination. However, the high-reliability sub-50-cm precision demanded by automated vehicles for lane-keeping and other applications, especially in urban areas, has significantly changed this landscape [169]. In most automated vehicles being developed, the GPS/GNSS receiver is relegated to a secondary sensor whose role is to loosely constrain (within a few meters) the primary localization sensors, usually camera(s) and/or lidar, to a global reference frame when building a digital map. The vehicles then locate themselves to decimeter accuracy within this digital map.

Automated driving does not necessarily demand sub-50-cm agreement between the coordinates of a given point in the digital map and the coordinates of the same point in a well-defined global reference frame. Rather, local self-consistency and accurate localization within the digital map is of greater importance. However, consistency of the digital map with a global coordinate frame is likely to become a pre-requisite for cooperative automated driving. If all collaborating vehicles navigate within the same digital map,

then precise exchange of information such as vehicle position, velocity, intent, etc. is possible [82, 83], even if the map itself is only globally accurate to a few meters. However, it is unlikely that automated vehicles from different manufacturers will rely on a common digital map. Consequently, the accuracy of the exchanged vehicle position is lower-bounded by the disagreement on the coordinates of the same physical location between different maps. Thus, exchange of accurate vehicle pose among vehicles, as well as other associated high-level information such as sensor data in the vehicle's body frame, will demand consistency among, or translation between, different digital maps.

Standard code-phase-based GNSS position measurements, such as those provided by all mass-market GNSS receivers, may be biased by as much as 3–5 meters on any given mapping session. Maps anchored by these measurements may not exhibit lane-level consistency with each other. One possible solution is to create digital maps with decimeter-accurate CDGNSS systems [64]. However, at current prices, such systems can only be installed on a limited fleet of specialized mapping vehicles. Precise point positioning (PPP) techniques offer a low-cost alternative to CDGNSS, but the frequent cycle-slipping experienced in urban areas impedes the convergence of PPP techniques [68].

This chapter explores the accuracy limit of globally-referenced mapping involving collaborating consumer vehicles whose sense of global position is based on standard code-phase-based GNSS receivers. Key parameters in this exploration are the asymptotic averages of the error sources that impair code-phase-based GNSS positioning: receiver thermal noise, satellite clock and orbit

errors, ionospheric and tropospheric modeling errors, and multipath. One or more vehicles navigating through a digital map over time make multiple time-separated GNSS measurements of the same location. If these vehicles collaboratively update the map over multiple sessions, then the GNSS errors are averaged across all sessions with appropriate weighting.

Are the GNSS errors at every map location—including deep urban locations—asymptotically zero-mean, or, on the contrary, do location-dependent biases persist in averages of time-separated standard GNSS measurements? Such is the question this chapter seeks to address.

5.3 Related Work

Improving the accuracy of maps by averaging GPS/GNSS tracks has been explored previously using a variety of approaches. An early effort, detailed in [128], proposed the precise determination of lane centerlines by clustering and averaging the GNSS tracks of probe vehicles. The accuracy of the estimated centerline was assessed in terms of the spread of GNSS tracks, assuming, without analysis, that the error was zero-mean at every location. More recently, [73] proposed vehicle lane determination via PPP on a rural road under open-sky conditions. The current chapter aims to perform localization at a similar accuracy level, but in urban and suburban areas and with the aid of a digital mapping sensor.

Minimizing the difference between GNSS measurements and the assigned map coordinates of locations visited multiple times by probe vehicles

has been a common feature of the seminal works on map-based precise localization in urban environments for automated driving [82, 83], but no analysis of the accuracy of the resulting map in the global coordinate system was provided.

The effect of multipath on measured pseudoranges was studied extensively for various signal types in [40]. However, this study was done under open-sky conditions with a static survey-grade antenna, hardly representative of a mass-market receiver in an urban environment. A detailed study on the distribution of code-phase and Doppler offsets of the multipath components from individual satellites in a dynamic urban setting was carried out in [170]. However, the error was characterized as the combined distribution of code phase delays over the entire duration of the run, which marginalizes over the temporally- and physically-local biases. On the contrary, this chapter explores the errors in the position domain for repeated sessions through a given realization of an urban corridor.

Other GNSS error sources such as errors in modeling of ionospheric [130] and tropospheric [23] delay have been studied extensively over many decades, and their long-term error characteristics have also been reported in the literature. However, the impact of these errors on the asymptotic statistics of code-phase-based GNSS position estimates has not been previously presented.

To the authors' best knowledge, despite the apparent simplicity of the problem, no prior work has studied the long-term statistics of GNSS errors in an urban environment representative of the conditions to be encountered by

consumer vehicles creating digital maps.

5.4 GNSS Error Analysis

5.4.1 Low-Cost GNSS in Urban Areas

Low-cost multi-GNSS receiver manufacturers have recently announced the development and release of low-cost multi-frequency multi-GNSS receivers [157]. Accordingly, the analysis in this section considers a vehicular platform equipped with a multi-frequency multi-GNSS receiver capable of tracking both code and carrier phase of GNSS signals.

Development of an extensive dense reference network in support of CDGNSS consumer vehicular positioning in urban areas, as suggested in [105], could be an expensive affair. PPP is a low-cost alternative to CDGNSS that requires only a sparse network of reference stations across the globe, but is not considered a viable option for urban GNSS positioning in this chapter because the constant cycle slips and outages experienced in urban areas [64] make it difficult for PPP's float carrier phase ambiguity estimates to converge [68], in which case PPP degrades to code-phase positioning accuracy.

While convergence of PPP carrier-phase ambiguities may be infeasible in urban areas, a partial PPP solution that exploits precise satellite orbits and clocks, as well as ionospheric and tropospheric corrections, can certainly improve the accuracy of code-phase-based GNSS position estimates. Since connected and automated vehicles will perforce enjoy network connectivity, this chapter assumes the availability of such GNSS corrections. Thus, the

kind of GNSS errors assessed in this section lie between those corresponding to the two extremes of standard standalone code-phase positioning and PPP. This type of GNSS positioning, hereafter referred to as enhanced code-phase positioning, exploits both code and carrier phase or frequency tracking, but, as opposed to PPP, does not attempt to estimate a quasi-constant float carrier phase ambiguity, making it suitable for urban applications.

5.4.2 Pseudorange Measurement

The pseudorange measurement at receiver $\mathtt{R}$ from satellite $\mathtt{S}_i$ is modeled as

$$\begin{aligned}\rho_i(t_{\mathtt{R}}) &= h_i[\boldsymbol{x}(t_{\mathtt{R}}), I_{\rho_i}(t_{\mathtt{R}}), T_i(t_{\mathtt{R}}), t_{\mathtt{R}}] + w_{\rho_i}(t_{\mathtt{R}}) \\ &= \Delta r_i + c[\delta t_{\mathtt{R}}(t_{\mathtt{R}}) - \delta t_{\mathtt{S}_i}(t - \delta t_{\mathrm{TOF}_i})] + I_{\rho_i}(t_{\mathtt{R}}) + T_i(t_{\mathtt{R}}) + w_{\rho_i}(t_{\mathtt{R}}), \quad (5.1)\end{aligned}$$

where

$$\boldsymbol{x}(t_{\mathtt{R}}) \triangleq \begin{bmatrix} \boldsymbol{r}_{\mathtt{R}}(t_{\mathtt{R}}) \\ \delta t_{\mathtt{R}}(t_{\mathtt{R}}) \end{bmatrix}$$

is the state of the receiver, comprising the receiver position, $\boldsymbol{r}_{\mathtt{R}}(t_{\mathtt{R}})$, at the time of the signal receipt event, $t_{\mathtt{R}}$, and the receiver clock bias, $\delta t_{\mathtt{R}}(t_{\mathtt{R}}) = t_{\mathtt{R}} - t$, with respect to true time t. The nonlinear measurement model is denoted by h_i; ρ_i denotes the measured pseudorange to $\mathtt{S}_i$; c denotes the speed of light in vacuum; $\delta t_{\mathtt{S}_i}(t) = t_{\mathtt{S}_i} - t$ denotes the satellite clock bias with respect to t; $\delta t_{\mathrm{TOF}_i}$ denotes the time-of-flight of the signal from $\mathtt{S}_i$, as an increment in true time; I_{ρ_i} and T_i denote the ionospheric and tropospheric delay experienced by the signal from $\mathtt{S}_i$, respectively; $w_{\rho_i} \sim (\mu_{w_i}, \sigma^2_{w_i})$ denotes

the sum of measurement thermal noise, multipath interference, non-line-of-sight (NLOS) delay, and other unmodeled errors; and Δr_i denotes the true range between $\mathtt{R}$ and $\mathtt{S}_i$, given as

$$\Delta r_i = \|\boldsymbol{r}_{\mathtt{R}}(t_{\mathtt{R}}) - \boldsymbol{r}_{\mathtt{S}_i}(t_{\mathtt{R}} - \delta t_{\mathtt{R}}(t_{\mathtt{R}}) - \delta t_{\mathrm{TOF}_i})\|,$$

where $\boldsymbol{r}_{\mathtt{S}_i}$ is the satellite position at the signal transmit event. Note from (5.1) that the receiver clock bias component of the state contributes identically to all pseudorange measurements.

Taking n_z pseudorange measurements $\{\rho_i\}_{i=1}^{n_z}$ and predictions $\bar{I}_{\rho_i}$ and $\bar{T}_i$ for each measurement, $\mathtt{R}$ estimates its state by solving a nonlinear least squares problem based on (5.1). First, it linearizes the measurement model in (5.1) about an initial guess of its state $\bar{\boldsymbol{x}}(t_{\mathtt{R}}) = \left[\bar{\boldsymbol{r}}_{\mathtt{R}}^T(t_{\mathtt{R}})\ \bar{\delta t}_{\mathtt{R}}(t_{\mathtt{R}})\right]^T$ and the modeled atmospheric delays:

$$\rho_i \approx h_i(\bar{\boldsymbol{x}}, \bar{I}_{\rho_i}, \bar{T}_i, t_{\mathtt{R}}) + \underbrace{\left[\frac{\partial h_i}{\partial \boldsymbol{x}}\right]^T_{\boldsymbol{x}=\bar{\boldsymbol{x}}}}_{H_i}(\boldsymbol{x} - \bar{\boldsymbol{x}}) + \tilde{I}_{\rho_i} + \tilde{T}_i + w_{\rho_i},$$

with $\tilde{I}_{\rho_i} = I_{\rho_i} - \bar{I}_{\rho_i}$ and $\tilde{T}_i = T_i - \bar{T}_i$. Representing all n_z measurements in matrix form yields

$$\begin{bmatrix} \rho_1 \\ \vdots \\ \rho_{n_z} \end{bmatrix} = \begin{bmatrix} h_1(\bar{\boldsymbol{x}}, \bar{I}_{\rho_1}, \bar{T}_1, t_{\mathtt{R}}) \\ \vdots \\ h_{n_z}(\bar{\boldsymbol{x}}, \bar{I}_{\rho_{n_z}}, \bar{T}_{n_z}, t_{\mathtt{R}}) \end{bmatrix} + \begin{bmatrix} H_1 \\ \vdots \\ H_{n_z} \end{bmatrix}(\boldsymbol{x} - \bar{\boldsymbol{x}}) + \begin{bmatrix} \tilde{I}_{\rho_1} \\ \vdots \\ \tilde{I}_{\rho_{n_z}} \end{bmatrix} + \begin{bmatrix} \tilde{T}_1 \\ \vdots \\ \tilde{T}_{n_z} \end{bmatrix} + \begin{bmatrix} w_1 \\ \vdots \\ w_{n_z} \end{bmatrix}$$

or

$$\boldsymbol{\rho} = \boldsymbol{h}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{I}}, \bar{\boldsymbol{T}}, t_{\mathtt{R}}) + H(\boldsymbol{x} - \bar{\boldsymbol{x}}) + \tilde{\boldsymbol{I}} + \tilde{\boldsymbol{T}} + \boldsymbol{w}. \tag{5.2}$$

Rearranging measured and modeled quantities on the left-hand side to get the standard form for a linearized measurement model yields

$$\boldsymbol{z} \triangleq \boldsymbol{\rho} - \boldsymbol{h}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{I}}, \bar{\boldsymbol{T}}, t_{\mathtt{R}}) + H\bar{\boldsymbol{x}} = H\boldsymbol{x} + \tilde{\boldsymbol{I}} + \tilde{\boldsymbol{T}} + \boldsymbol{w}. \tag{5.3}$$

The ith row of the measurement sensitivity matrix H is

$$H_i \approx \left[\frac{\bar{\boldsymbol{r}}_{\mathtt{R}}^T(t_{\mathtt{R}}) - \bar{\boldsymbol{r}}_{\mathtt{S}_i}^T(t_{\mathtt{R}} - \bar{\delta}t_{\mathtt{R}} - \bar{\delta}t_{\mathrm{TOF}_i})}{\|\bar{\boldsymbol{r}}_{\mathtt{R}}(t_{\mathtt{R}}) - \bar{\boldsymbol{r}}_{\mathtt{S}_i}(t_{\mathtt{R}} - \bar{\delta}t_{\mathtt{R}} - \bar{\delta}t_{\mathrm{TOF}_i})\|}, \quad 1 \right].$$

By solving (5.3) for $\boldsymbol{x}$, updating $\bar{\boldsymbol{x}}$, and iterating until convergence, R obtains its state estimate $\hat{\boldsymbol{x}}(t_{\mathtt{R}})$:

$$\hat{\boldsymbol{x}}(t_{\mathtt{R}}) \triangleq \begin{bmatrix} \hat{\boldsymbol{r}}_{\mathtt{R}}(t_{\mathtt{R}}) \\ \hat{\delta}t_{\mathtt{R}}(t_{\mathtt{R}}) \end{bmatrix}.$$

For dynamic applications such as vehicle tracking, the state $\boldsymbol{x}(t_{\mathtt{R}})$ is typically augmented to include the time derivatives of $\boldsymbol{r}_{\mathtt{R}}(t_{\mathtt{R}})$ and $\boldsymbol{t}_{\mathtt{R}}(t_{\mathtt{R}})$, and the measurement model typically assumes direct measurement of apparent Doppler frequency.

5.4.3 Error Sources

The major sources of error in the estimates $\hat{\boldsymbol{r}}_{\mathtt{R}}$ and $\hat{\delta}t_{\mathtt{R}}$ are as follows:

5.4.3.1 Thermal Noise

Measurement thermal noise at the receiver is one of the components of w_{ρ_i} in (5.1). The effect of thermal noise can be accurately modeled as a white Gaussian random variable with zero mean and standard deviation σ_{T}. For

the pseudorange measurement, σ_{T} is typically between 10–30 cm, depending on the signal carrier-to-noise ratio, signal bandwidth, and receiver tracking bandwidth [96]. Estimation of the receiver state from multiple appropriately-weighted measurements with independent thermal-noise errors, and processing such measurements over time through a filter based on the modeled dynamics of the receiver, renders negligible the position-domain effects of uncorrelated zero-mean thermal noise. As a result, thermal noise is not a major contributor to the asymptotic accuracy of a digital map.

5.4.3.2 Satellite Orbit and Clock Errors

Satellite orbit and clock errors manifest in the modeled satellite position $\bar{r}_{\mathtt{s}_i}$ and the modeled satellite clock bias $\bar{\delta t}_{\mathtt{s}_i}$. The International GNSS Service (IGS) provides orbit and clock models for GNSS satellites. The predicted *ultra rapid* orbits and satellite clocks have an accuracy of ∼5 cm and ∼3 ns, respectively [1]. These may add up to ∼1 m of combined pseudorange model error for a given satellite. The 17-h retroactively-available *rapid* orbits and satellite clock models are accurate to ∼2.5 cm and ∼75 ps RMS errors, respectively [1], adding up to less than 5 cm of RMS error in the modeled pseudorange for a given signal. Since the orbit and clock parameters are fit to measurements made at IGS analysis centers, the errors in the estimated parameters must be asymptotically zero-mean by design of the estimator. For post-processing applications such as mapping, it is reasonable to assume the availability of *rapid* orbit and satellite clock products, and thus the asymptotic average position

errors due to errors in modeled satellite position and clock bias can be reduced to a sub-5-cm level.

5.4.3.3 Ionospheric Modeling Errors

The code-modulated GNSS signal propagates slower through the ionosphere as compared to vacuum due to the slightly-greater-than-unity *group* index of refraction for this atmospheric layer. The excess group delay is given as

$$\Delta\tau_g = \frac{40.3 \cdot TEC}{cf^2},$$

where TEC is the total electron content in electrons/m^2 and f is the frequency of the propagating signal. At the GPS L1 frequency centered at 1575.42 MHz, the excess ionospheric group delay is roughly 16.24 cm per TECU (1 TECU $\triangleq$ 10^{16} electrons/m^2). If not modeled, the ionospheric delay can lead to ranging errors greater than 15 m.

The ionospheric delay can be estimated via an ionosphere model or, in case of a multi-frequency receiver, eliminated via a combination of multiple-frequency pseudorange measurements. The latter technique does not require any external aiding, but the formation of the ionosphere-free combination exacerbates pseudorange noise, including any biases due to tracking of multipath signals. Compensating for ionospheric delay with the aid of an ionosphere model is applicable to both single- and multi-frequency receivers. It relies on accurate delay modeling based on ionospheric measurements at permanent GNSS reference stations, such as those that form the IGS network. While

both methods have their merits, the analysis in this section considers corrections from an ionospheric model, and thus will not be relevant to applications where the ionosphere-free combination is applied. Note that those applications would likely experience worse multipath errors than the ones presented later, requiring a separate multipath analysis along the lines of Sec. 5.4.3.5.

Ionospheric model accuracy was studied comprehensively in [130]. The method in [130] generates unambiguous carrier-phase measurements from a global distribution of permanent receivers to compute the true slant total electronic content (STEC) for each satellite, and compares the model prediction for a number of models with the ground truth. In [129], the same authors compared PPP convergence times when applying different ionospheric correction models. This section extends the analysis in [130] and [129] to examine whether there exist long-term position-domain biases in enhanced code-phase positioning.

The post-fit residuals for multiple regional and global ionospheric models, computed as described in [130], were graciously made available by the authors of [130] for the year 2014. These residuals were computed for GPS signals as observed at about 150 reference stations around the globe at 5 min intervals.

To observe the position-domain effect of the ionospheric modeling errors in isolation, this section neglects all other error sources, reducing the linearized

measurement model in (5.3) to

$$\boldsymbol{z} = H\boldsymbol{x} + \tilde{\boldsymbol{I}}.$$

Historical GPS satellite almanacs can be combined with the timestamps from the residuals data to obtain the measurement sensitivity matrix H at each epoch for each station. With an elevation-dependent measurement covariance matrix R, the error in the weighted least-squares solution due to errors in ionospheric modeling is

$$\hat{\boldsymbol{x}} - \boldsymbol{x} = \left(H^T R^{-1} H\right)^{-1} H^T R^{-1} \tilde{\boldsymbol{I}}.$$

Fig. 5.1 summarizes the results for ionospheric corrections obtained from the IGS global ionospheric map (GIM). Each of the arrows in Fig. 5.1 points in the direction of the position bias in the east-north plane, as estimated over 12 months of data from 2014 (more than 800,000 samples per station). The magnitude of the horizontal position bias is depicted by the color of the arrow according to the scale shown on the right. Interestingly, there is a clear trend of southward bias in the position error for most stations in the northern hemisphere, and a mild trend of northward bias in the position error for stations in the southern hemisphere. A numerical summary of the IGS GIM position bias is presented in Table 5.1, along with a similar analysis for the Wide Area Augmentation System (WAAS) ionospheric corrections available for the contiguous United States (CONUS) region. As reported in [130], the

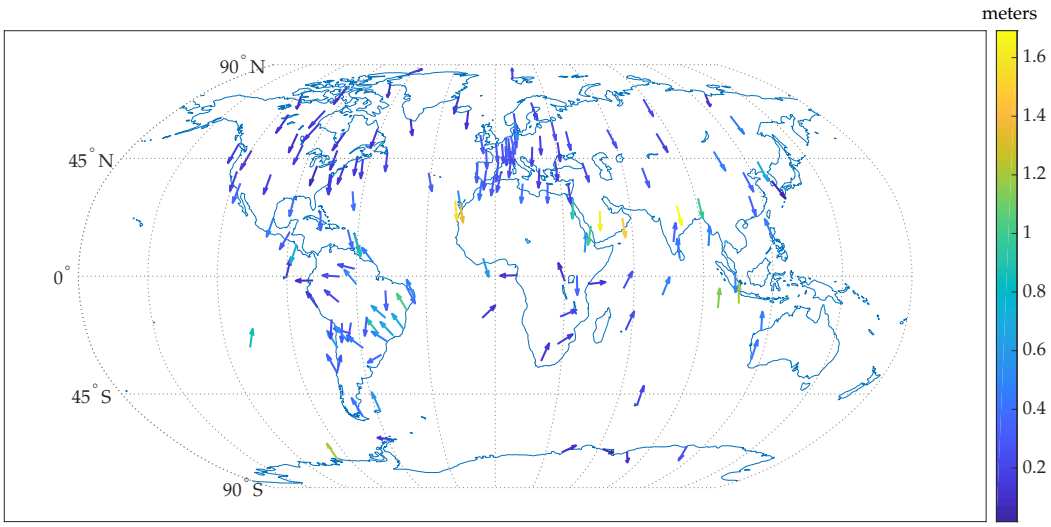


Figure 5.1: Direction and magnitude (the latter represented by color, in meters) of the long-term average horizontal position error due to errors in the delay estimates provided by the IGS GIM. Note that the meridians are curved outwards due to projection of the spherical map, and that arrows parallel to the curved meridians point directly south or north.

WAAS model was found to exhibit a significantly smaller RMS error in ionosphere TEC estimates when compared to the IGS GIM; however the long-term position bias due to WAAS corrections is similar to or worse than those for the IGS model.

Another global ionospheric model, the Fast PPP model [129], was also studied as above. Fast PPP natively models the ionosphere as a two-layered shell, but is also made available in the standard one-layer IONEX (ionosphere-map exchange) format [130] for dissemination. The results presented in Table 5.1 represent the IONEX version of Fast PPP. In comparison with the IGS corrections, it is clear that the Fast PPP IONEX GIM corrections result in

Table 5.1: Long-term average position error due to ionospheric model errors (ϕ denotes station latitude). IGS: International GNSS Service; PPP: precise point positioning; WAAS: Wide Area Augmentation System; CONUS: contiguous United States; IONEX: ionosphere-map exchange format

Ionosphere Model	**Region**	**East (m)**	**North (m)**	**Up (m)**
IGS	$\phi \geq 30^\circ$	0.0107	−0.2129	0.6733
	$30^\circ > \phi > -30^\circ$	−0.0651	−0.0692	1.5467
	$\phi \leq -30^\circ$	0.0237	0.2450	0.3355
WAAS	CONUS	−0.0048	−0.2916	−0.1248
Fast PPP IONEX	$\phi \geq 30^\circ$	−0.0042	−0.0099	−0.0122
	$30^\circ > \phi >$-30°	−0.0390	0.0013	−0.3053
	$\phi \leq$-30°	−0.0325	−0.0087	0.0309

substantially unbiased long-term position errors at the global test locations. However, it must be conceded that the results in Table 5.1 are best-case results, as they are based on data from the same permanent reference stations used to constrain the model.

To understand the reason behind the systematic biases with IGS corrections, note that any ionospheric modeling bias that identically affects all satellites does not have any impact on the accuracy of the GNSS position solution, as this common error is absorbed in $\hat{\delta}t_{\mathtt{R}}$. Rather, position-domain biases arise from the azimuthal- and elevation-dependence of ionosphere model errors. From analysis of the spatial distribution of post-fit residuals, it was found that appreciable azimuthal and elevation residual gradients persist in the IGS ionospheric corrections. These gradients are represented graphically in Fig. 5.2 for one representative station from the northern hemisphere (station code: EUSK,

latitude: 50°40′26.87″, longitude: 6°45′48.72″) and one representative station from the southern hemisphere (station code: VACS, latitude: -20°17′48.47″, longitude: 57°29′13.79″). The post-fit residuals are binned in azimuth and elevation and the average value in each bin is denoted by the color of the representing disc. The size of the disc denotes the number of samples of post-fit residuals available in each bin. Due to the inclination angle of the GPS satellite orbits, the angular distribution of satellites at any given latitude is non-uniform.

From Fig. 5.2, it is clear that the elevation gradients in the ionospheric residuals are pronounced. A subtle azimuthal gradient also exists, mainly along the north-south direction. Such spatial non-uniformity, coupled with the non-uniform satellite angular distribution, may be the reason for the observed persistent position biases. While the elevation gradients are consistent for stations at all locations, the azimuthal gradients appear to invert along the north-south direction between the northern and southern hemisphere. This is likely the reason for the opposite direction of the average horizontal position bias in the northern and southern hemispheres.

Such persistent position-domain biases due to inaccurate ionospheric modeling have not been previously reported in the literature, and are a rather remarkable result. While some single-frequency PPP (SF-PPP) techniques eliminate the ionospheric delays based on the GRAPHIC combination [175], many other techniques that rely solely on ionospheric corrections from GIMs have been shown to achieve 30 cm 95% accuracy in the east-north plane after

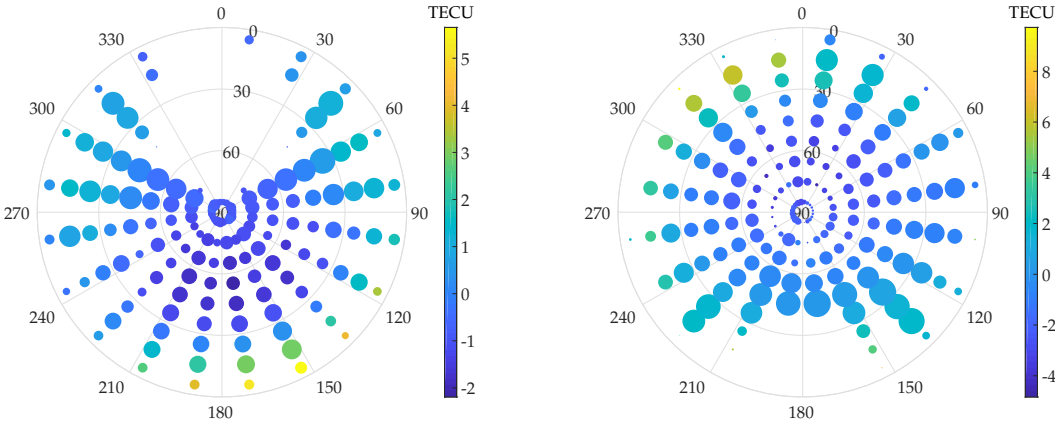


Figure 5.2: Azimuth and elevation dependence of post-fit IGS global ionospheric map (GIM) residuals. (**a**) A representative station from the northern hemisphere. (**b**) A representative station from the southern hemisphere. The average residual error (in TECU) is denoted by the color of the disc. The size of the disc indicates the number of samples of post-fit residuals available in each bin.

convergence, with sub-10-cm bias [160], seemingly contradicting the results here. The key difference is that the SF-PPP methods involve estimation of a float carrier ambiguity term for each satellite arc. A portion of the systematic biases in the GIM estimates is likely absorbed in these states of the estimator, thereby attenuating the position biases in the east-north plane. For instance, the SF-PPP technique in [160] is based on the phase-adjusted pseudorange algorithm [78], wherein the ambiguity term for each satellite, physically an unknown constant, is in fact iteratively estimated with small but non-zero process noise. In such an estimator, the ambiguity term can absorb slowly time-varying systematic biases. In other SF-PPP techniques, the ionospheric correction term is explicitly included as a state to be estimated, and

the estimates from GIM are applied as pseudo-observations [116, 143]. Once again, decimeter-level biases in the GIM estimates of the ionospheric delay may not necessarily appear in the final reported position accuracy of the SF-PPP method. Of course, such absorption of biases in augmented states is not undesirable. However, for the case of urban vehicular positioning, convergence of SF-PPP is a concern due to carrier phase cycle slipping, as discussed earlier. In an enhanced code-phase-based receiver, the high variance of the code noise leads to poor observability of the decimeter-level horizontal position bias due to ionospheric modeling errors. Thus, ionospheric biases are not often estimated in a code-phase-based GNSS estimator.

Another factor of note is that 2014 was a maximum in the 11-year solar activity cycle, and thus the IGS GIM accuracy may have been worse than usual over this period of time.

In conclusion, persistent decimeter-level biases in the east-north plane and meter-level biases in the vertical direction can arise when ionospheric delay corrections are sourced from the IGS GIM, or similar, even under ideal open-sky conditions. More advanced models of the ionosphere with more accurate slant TEC measurements may achieve better results. Elimination of the ionospheric delay based on the ionosphere-free combination is another option, but tends to worsen multipath-induced position errors. If corrections from some ionosphere model lead to unbiased position errors, then for globally-referencing digital maps by averaging GNSS measurements over many sessions it is advisable to avoid the combination of multi-frequency signals.

5.4.3.4 Tropospheric Modeling Errors

In the troposphere, or more generally the neutral atmosphere, the index of refraction departs from unity much less than in ionosphere at GNSS frequencies, causing a delay of ∼2.4 m at zenith. The index of refraction in the troposphere is non-dispersive, and thus cannot be estimated using multiple-frequency signals. The tropospheric delay is obtained from models of the climatological parameters (temperature, pressure, and water vapor pressure) along the propagation path.

State-of-the-art tropospheric models [23] fit a small number of location- and day-of-year-dependent coefficients to climatological data from numerical weather models (NWMs) to model the zenith tropospheric delay. The zenith delay is mapped to any elevation angle using mapping functions [22]. Similar to the ionospheric models, the tropospheric mapping functions may introduce a differential azimuth- and elevation-dependent error. For empirically-derived mapping functions such as VMF1 [22] and GMF [24], the mean error at lowest elevation of 5° has been shown to be under 50 mm (this value is typically reported as 10 mm station height error, which is approximately one-fifth of the delay error at lowest elevation [22]). As a result, this chapter assumes that time-averaged tropospheric model errors would introduce sub-5-cm errors in the position domain, and would thus not impede asymptotically accurate collaborative mapping in both horizontal and vertical components at the several-decimeters level.

5.4.3.5 Multipath Error

In ideal circumstances, each signal received from an overhead satellite arrives only along the least-time path. In practice, however, this so-called line-of-sight (LOS) component is accompanied by other components due to signal diffraction and single- or multiple-signal reflections off surrounding surfaces and obstacles (e.g., the glass facade of a nearby building, poles, trees, etc.). The complex baseband representation of the N signal components received from a particular satellite at a particular frequency and code is

$$r(t) = \sum_{i=0}^{N-1} A_i(t) C[t - \tau_i(t)] \exp[j\theta_i(t)],$$

where A_i is the amplitude of the ith component, $C(t)$ is the GNSS code modulation, $\tau_i(t)$ is the delay of the ith signal component relative to an unobstructed LOS signal, and $\theta_i(t)$ is the beat carrier phase of the ith component. The combination of multiple components distorts the received signal and causes errors in the pseudorange and phase measurements.

Unlike the study of ionospheric modeling errors, for application in urban mapping, multipath errors cannot be characterized with data from survey stations with a clear view of the sky. This section considers a simulation approach for scalable analysis of multipath tracking errors in an urban environment. The objective of this study was to inspect the presence of persistent biases caused by multipath due to the surrounding structure in the navigation solution averaged over multiple sessions

Table 5.2: Urban scenario parameters used in the multipath simulation.

Distance from road center to buildings	24 m
Distance from road center to vehicle	5 m
Mean distance between road center and trees	20 m
Antenna height	2 m
Mean building width	30 m
Building width standard deviation	25 m
Mean building height	40 m
Building height standard deviation	20 m
Probability of gap between buildings	0.5
Mean gap width	30 m
Mean distance between trees	60 m
Mean distance between poles	25 m

Scenario Setup The present simulation study was based on the open-access Land Mobile Satellite Channel Model (LMSCM) [79], itself based on extensive experimentation with a wideband airborne transmitter at GNSS frequencies in urban and suburban environments. First, an urban corridor was simulated stochastically following the procedure described in [80]. The corridor was composed of buildings, trees, and poles. Some of the important parameters for the generation of the scene are summarized in Table 5.2, and a part of the scene realization is shown in Fig. 5.3. Multi-GNSS satellite trajectories were generated at randomly-selected times based on GPS and Galileo satellite almanac data. An average of 25 satellites were available above an elevation mask of 5°, consistent with modern multi-GNSS receivers. The satellites were assumed to be stationary over the simulation period of 60 s. Navigation solution errors were computed over 1000 60-s sessions.

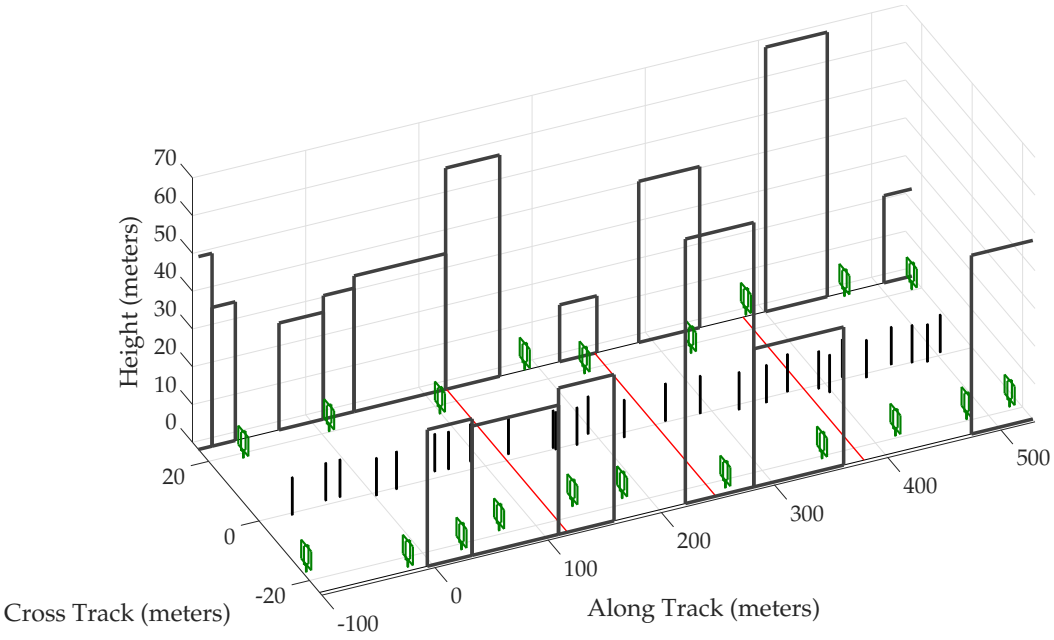


Figure 5.3: Initial segment of the simulated urban corridor. Red lines across the road denote the positions where the vehicle is momentarily stopped.

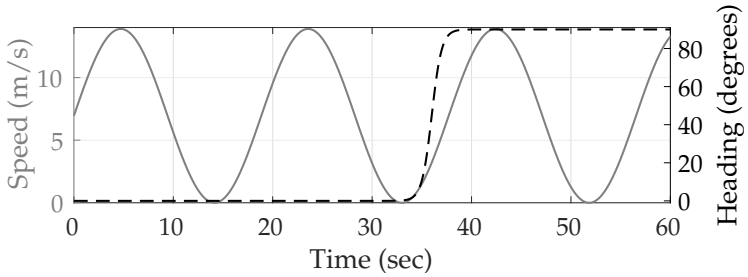


Figure 5.4: Vehicle speed (solid line) and heading (dashed line) simulating stop-and-go motion with a 90° right turn.

The vehicle trajectory was kept consistent across all 1000 driving sessions to avoid decorrelation of multipath error due to variable receiver motion. The trajectory was parametrized by its speed and heading, as described in [80]. The vehicle started at the zero coordinate on the along-track axis, and traveled in the positive direction, which was assumed to be aligned with the local north. The simulated trajectory was 60 s long and simulated a vehicle in stop-and-go traffic executing one 90° right turn, as shown in Fig. 5.4. The vehicle traveled roughly 430 m and faced eastwards at the end of the trajectory. The three low-speed intervals, marked with red line segments in Fig. 5.3, are expected to present severe multipath effects since multipath errors decorrelate slowly, and thus tend to reinforce one another within the navigation filter, when the vehicle moves slowly.

Multipath Simulation The LMSCM generates power, delay, and carrier phase for N LOS and echo signals. The interaction of the LOS with the simulated obstacles is governed by deterministic models for attenuation, diffraction,

and delay. The LOS components of the combined signal, denoted $r_{\mathrm{LOS}}(t)$, may be composed of multiple components due to signal diffraction. These components are modeled as

$$r_{\mathrm{LOS}}(t) = \sum_{i=0}^{N_{\mathrm{LOS}}-1} A_i(t)C[t-\tau_i(t)]\exp[j\theta_i(t)].$$

In the special case of an unobstructed LOS signal, $N_{\mathrm{LOS}} = 1$, $A_0(t) = 1$, $\tau_0(t) = 0$, and

$$\theta_0(t) = \frac{\|\boldsymbol{r}_{\mathtt{R}}(t) - \boldsymbol{r}_{\mathtt{S}}(t)\| \cdot 2\pi}{\lambda} + \gamma_0,$$

where λ denotes the wavelength and γ_0 is a constant due to phase initialization in the satellite and receiver [124].

The LMSCM generates the $N - N_{\mathrm{LOS}}$ NLOS echoes stochastically based on satellite azimuth and elevation, receiver dynamics, and general characteristics of the scene (e.g., an *urban car* scenario). This stochastic procedure might not be representative of multipath over multiple sessions through the same urban corridor, where certain echoes might persist over different sessions. To address this limitation, the LMSCM was augmented by the present authors to generate one- and two-bounce deterministic reflective NLOS echoes off the simulated buildings, and a one-bounce NLOS echo off the ground surface. These three additional reflective NLOS echoes, denoted $r_{\mathrm{DET}}(t)$, were added to $r(t)$ and are modeled as

$$r_{\mathrm{DET}}(t) = \sum_{i=N}^{N+2} b_i(t)A_i(t)C[t-\tau_i(t)]\exp[j(\theta_i(t)+\theta_i'(t))],$$

where $b_i(t) \in \{0, 1\}$ denotes whether the surrounding geometry supports the reflective echo. Since these reflections are expected to be the stronger than other diffracted and multiple-bounce NLOS echoes, the amplitudes $A_i(t), i \in \{N, N+2\}$ for reflective echoes were drawn from the distribution of the strongest echo generated stochastically by the LMSCM at each epoch. By experiment, this distribution was found to be log-normal with $20\log_{10}(A_i) \sim \mathcal{N}(-22, 5), i \in \{N, N+2\}$. The delays for the reflective echoes are given as

$$\tau_i(t) = \frac{\|\boldsymbol{r}'_{\mathtt{R}}(t) - \boldsymbol{r}_{\mathtt{S}}(t)\| - \|\boldsymbol{r}_{\mathtt{R}}(t) - \boldsymbol{r}_{\mathtt{S}}(t)\|}{c},\ i \in \{N, N+2\},$$

where $\boldsymbol{r}'_{\mathtt{R}}(t)$ is the position of the imaginary *image* antenna [26] about the reflecting plane (building or ground). Similarly, the carrier-phase of the reflective echoes is computed geometrically as

$$\theta_i(t) = \frac{\|\boldsymbol{r}'_{\mathtt{R}}(t) - \boldsymbol{r}_{\mathtt{S}}(t)\| \cdot 2\pi}{\lambda} + \gamma_0,\ i \in \{N, N+2\}.$$

A random carrier-phase offset $\theta'_i(t) \in [0, 2\pi)$ was added at the reflection point every time a new reflective echo was spawned to simulate the material-specific phase offset introduced by the reflection process.

Receiver A receiver simulator was developed to account for the mediating effects that a receiver's tracking loops and navigation filter have on multipath-induced position errors in a receiver's reported position solution. The simulated receiver tracks the combination of all N_{LOS} line-of-sight signals and $N + 2 - N_{\text{LOS}}$ multipath echoes for a given signal. If $R(\tau)$ denotes the correlation function of the GNSS signal's spreading code, then the multipath delay

error in the tracked code phase, relative to unobstructed LOS, is given as the solution to [26]

$$0 = S_{\text{coh}}(\tau) = \sum_{i=0}^{N+2} A_i \cos(\theta_i - \theta_c) \times \left[R\left(\tau - \tau_i + \frac{d}{2}\right) - R\left(\tau - \tau_i - \frac{d}{2}\right) \right],$$

where θ_c is the tracked carrier-phase of the combined signal:

$$\theta_c = \text{atan2}\left(\sum_{i=0}^{N+2} A_i R(\tau_c - \tau_i) \sin(\theta_i), \sum_{i=0}^{N+2} A_i R(\tau_c - \tau_i) \cos(\theta_i) \right).$$

The parameter d is the early-to-late correlator spacing in the receiver. It is well-known that a wide-bandwidth receiver with narrow correlator spacing mitigates the effect of multipath [26]. To this end, the receiver considered in this simulation implements $d = 0.1$. It must be mentioned that $R(\tau)$ was implemented as the correlation function for GPS L1 C/A identically for all the simulated signals. Modernized GNSS signals have better multipath mitigation characteristics [40], but this behavior was not included in the simulation.

Another important observation is that when the LOS signal is strong as compared to the echo signals, the time derivative of the tracked carrier-phase is equal to the Doppler frequency of the LOS signal, which changes smoothly in accordance with the motion between the satellite and the receiver. However, when the LOS signal is comparable to or weaker than other rapidly-decorrelating echoes, the combined carrier-phase is uniformly random. In a GNSS receiver, the phase lock loop's phase-lock indicator indicates whether a sufficiently strong LOS signal is available, enabling carrier lock [64]. The simulator's phase-lock indicator is asserted only if (1) the tracked Doppler

frequency does not deviate significantly from a second-order polynomial, and (2) the strongest received component (either LOS or NLOS) is attenuated no more than 25 dB with respect to an unattenuated signal.

Navigation Filter At each epoch, n_z multipath-free, ionosphere-free, and troposphere-free simulated pseudorange measurements were combined with corresponding simulated multipath tracking delay errors and fed to a navigation filter that estimates the receiver state. The navigation filter implemented in this chapter is an extended Kalman filter (EKF) with a nearly constant velocity motion model following [11]. The standard details of the EKF are omitted for brevity.

The effect of multipath tracking on the navigation solution is strongly dependent on the receiver's multipath rejection scheme. Two schemes are explored here. The first is a hypothetical ideal multipath rejection scheme that excludes all signals for which the LOS signal has a smaller-than-10-dB advantage over its multipath echoes. The second scheme implements a normalized innovation squared (NIS) test to reject multipath signals based on measurement innovations [11]. At the $(k+1)$th measurement update step, the difference between the predicted and observed measurement vector, called the innovation and denoted $\boldsymbol{\nu}(k+1)$, is squared and normalized by its covariance, which is the sum of the measurement covariance matrix, $R(k+1)$, and the propagated state covariance transformed through the measurement sensitivity matrix, $H(k+1)P(k+1|k)H(k+1)^T$. In the absence of multipath tracking

errors, the resulting NIS statistic is chi-squared distributed with n_z degrees of freedom. If the NIS statistic exceeds a chosen threshold, then the signal with the largest normalized innovation is dropped. This continues until the NIS statistic falls below the threshold or the number of remaining signals drops to a preset minimum number of required signals.

Simulation Results Fig. 5.5 shows the mean position error in the east, north, and up directions over 1000 sessions for the two multipath rejection schemes mentioned previously. From Fig. 5.5a, it can be seen that sub-20 cm average error is achievable with hypothetical ideal multipath exclusion. A closer look at Figs. 5.4 and 5.5a reveals that the decimeter-level sinusoidal position error trend, initially in the north direction and later in the east direction, in fact corresponds with the along-track accelerations of the vehicle that were not adequately tracked by the nearly-constant-velocity-model-based navigation filter.

Fig. 5.5b shows that the NIS test based exclusion of signals was able to approach the performance of ideal exclusion in the horizontal plane, save for the first stationary period where the vehicle was moving at low speed between buildings on both sides. The average vertical position error was much worse, growing as large as 1.75 m in magnitude.

To determine whether the average errors shown in Fig. 5.5 are in fact persistent biases, a study of the standard deviation of position errors was conducted. The standard deviation of the average errors in east, north, and

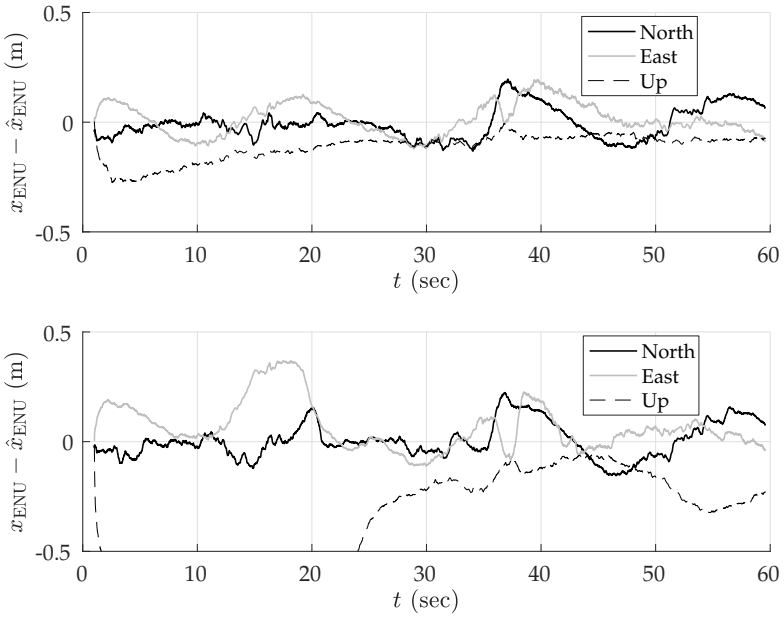


Figure 5.5: Mean position error in the east-north-up (ENU) frame over 1000 sessions due to multipath. (a) Ideal multipath exclusion. (b) NIS-based multipath exclusion. The black, gray, and dashed-black lines represent the error in the east, north, and up directions, respectively. The up error in the bottom panel reached a maximum magnitude of 1.75 m.

up directions was computed for disjoint averaging ensembles of size 1, 2, 4, 8, 16, 32, 50, and 100 sessions taken from the total of 1000 simulated sessions. For instance, 125 disjoint ensembles of eight sessions were selected, and the position errors were averaged over the eight sessions in each set. The standard deviation of the eight-session-averaged errors was then computed across the 125 ensembles. In the case of an averaging ensemble with only a single session (i.e., no averaging), the computed standard deviation is simply the measured standard deviation of the position error across all 1000 simulated runs. In the case of averaging over 100 sessions, the standard deviation is computed based on 10 disjoint averaging ensembles of 100 sessions each.

Note that because the simulation study was based on the same 1000 simulations for all averaging ensembles, the east, north, and up means taken across all averaging ensembles are equivalent to those shown in Fig. 5.5. The more interesting trend is the decreasing standard deviation with increasing size of the averaging ensemble, as shown in Fig. 5.6 for the case of NIS-based multipath rejection and for the east and north error components. As expected, the standard deviation of errors was higher at locations where the vehicle moved at low speed and multipath decorrelated slowly. Additionally, the standard deviation was larger at the beginning of the trajectory where the street was lined with tall buildings on both sides.

The standard deviation of the average east and north position error over 100 sessions was bounded below 15–20 cm. Thus, it is highly likely that the ∼40-cm error in the north direction between 15–20 s in Fig. 5.5b is in fact

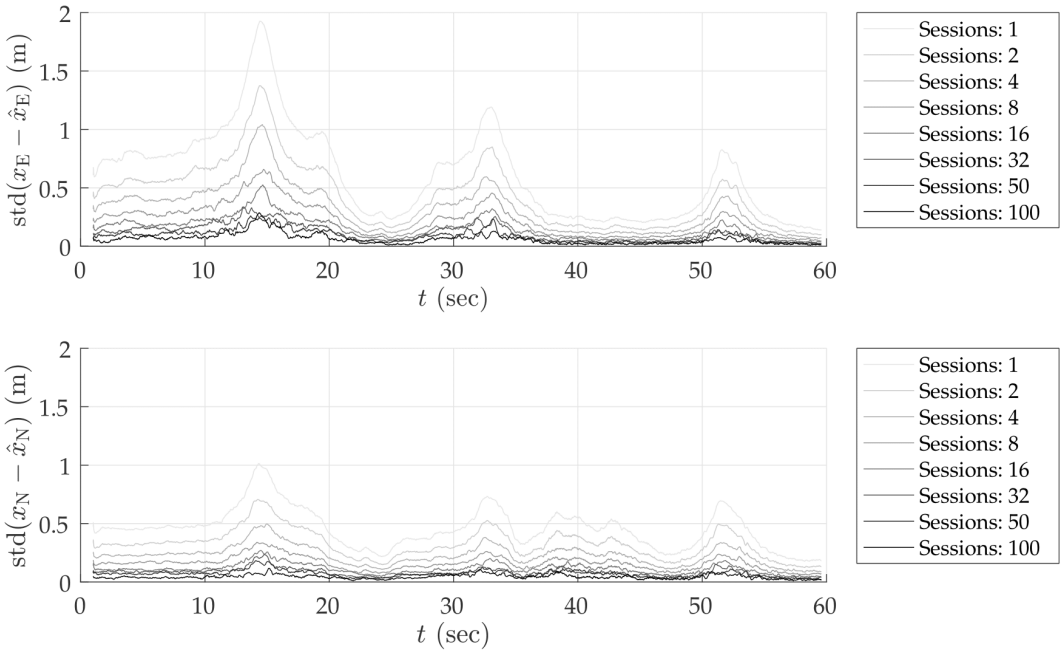


Figure 5.6: Standard deviation of average position error in east and north directions for NIS-based multipath exclusion as a function of the number of sessions over which the errors are averaged. Top panel: standard deviation in the east direction. Bottom panel: standard deviation in the north direction.

a persistent non-zero bias.

Table 5.3 summarizes the results of the multipath simulation study. It shows the 95-percentile horizontal error magnitude for increasing averaging ensemble sizes and for both ideal and NIS-based multipath exclusion. The 0–60 s average case lists the 95-percentile error over the entire trajectory, whereas the 13–19 s average case lists the 95-percentile error in the worst-case segment of the trajectory in terms of horizontal position bias and standard deviation. This challenging segment is illustrative of persistent problem spots that will arise in urban areas, within which multipath-induced biases will be larger than average. As expected, the 95-percentile error in Table 5.3 shrank as the averaging ensemble size became larger. For the urban corridor and vehicle dynamics considered in this simulation, NIS-based exclusion achieved 35 cm 95-percentile horizontal error with averaging over 100 sessions. Even in the worst-case region of the trajectory, the 95-percentile horizontal error remained below 50 cm. As multipath exclusion approaches the ideal case, with aid from other sensors or a 3D model of the surroundings, for example, the 95-percentile horizontal error could be reduced to as low as 25 cm for the simulated corridor.

5.5 Empirical Results

To validate the results obtained in the above analyses, GNSS and visual data were collected in a moderate urban area north of the University of Texas at Austin campus in Austin, TX. This section presents the error statistics of

Table 5.3: 95-percentile horizontal errors for increasing averaging ensemble sizes and for both ideal and NIS-based multipath exclusion.

Averaging Ensemble Size:		**1**	**2**	**4**	**8**	**16**	**32**	**50**	**100**
Ideal	0–60 s average (m)	1.5910	1.1262	0.7902	0.5488	0.4078	0.3090	0.2696	0.2147
	13–19 s average (m)	2.5925	1.7809	1.2136	0.8927	0.6416	0.4145	0.3544	0.2609
NIS	0–60 s average (m)	1.7851	1.2795	0.9245	0.6588	0.5169	0.4175	0.3920	0.3526
	13–19 s average (m)	3.1217	2.1953	1.5467	1.1720	0.8456	0.6470	0.5950	0.4702

various flavors of code-phase GNSS positioning.

5.5.1 Rover and Reference Platforms

The rover GNSS receiver is one among several sensors housed inside the Sensorium (described in Sec. 4.6). The GNSS data are processed by a software-defined GNSS receiver tracking signals from GPS L1 C/A, GPS L2CLM, Galileo E1, and SBAS. Data from both the primary (passenger's side) and secondary (driver's side) antennas are used to reconstruct a sub-dm-accurate CDGNSS-based ground truth trajectory, as described in [64]. Enhanced code-phase positioning is performed on the data from the primary antenna, incorporating precise orbit and clock products from IGS, ionospheric corrections from WAAS satellites, and the Saastamoinen model for tropospheric corrections, in addition to NIS-based exclusion of multipath signals. Double-differenced pseudorange-based positioning is also performed with the data from the primary antenna, as discussed later in this section. The code-phase-based position estimates are compared against the ground truth from the primary antenna to study the code-phase positioning error statistics. The primary antenna feed is also input to a ublox M8T receiver for comparison against the enhanced code-phase software receiver.

5.5.2 Test Route

The test route is a 1-km loop north of the University of Texas at Austin campus in Austin, TX. The route includes a variety of light-to-moderate urban

Figure 5.7: An overview of the 1-km test route. The Dean Keeton corridor, toward the left, is spanned by a pedestrian bridge and flanked by buildings on both sides. A total of 75 laps of the test route were driven over six separate campaigns.

conditions, from open-sky to overhanging trees to built-up areas. The Dean Keeton corridor, toward the left in Fig. 5.7, is the most challenging stretch along the test route for GNSS positioning. It passes below a pedestrian bridge and is flanked on both sides by buildings ranging from 30 to 65 meters tall set back 28 meters from the center of the roadway.

To study the code-phase-based positioning error characteristics over time-separated sessions in the same area, multiple laps of the test route were driven over six separate campaigns. The first two campaigns were conducted on 21 December 2017 and 15 January 2018, while the other four campaigns were conducted in pairs of two on 3 June 2018 and 4 June 2018. The GNSS error charts are presented for a total of 75 laps of the test route.

5.5.3 Empirical GNSS Error Analysis

Fig. 5.8 shows the error in the enhanced code-phase GNSS position solutions with respect to the ground truth. The error is plotted versus the distance along the 1-km loop. The beginning of this loop is taken to be immediately after the overhead pedestrian bridge along the Dean Keeton corridor. It is observed that the enhanced code-phase GNSS errors are clustered separately for each of the campaigns, and that each cluster is offset from zero by as much as 1 m in the horizontal plane. Such error characteristics are representative of ionospheric modeling errors, which have a long decorrelation time. It is also evident that the error variance is larger as the receiver exits the challenging portion of the loop at which point the tracking loops were recovering from signal loss under the bridge. The effect is especially pronounced in the vertical direction. Fig. 5.9 shows similar errors for the commercial ublox M8T receiver. The error traces from the ublox receiver show a wider spread than the enhanced code-phase receiver, likely due to lack of precise orbit and clock corrections.

On the basis of Figs. 5.8 and 5.9, one might be tempted to conclude that errors in enhanced code-phase and stand-alone GNSS navigation solutions are substantially non-zero-mean, especially in the north and up directions, despite the overhead GNSS constellation changing substantially between sessions. It certainly appears that the permanent structures (buildings, bridge) along the test loop left a bias in the vertical direction during the first 400 m along the loop. However, the bias in the north direction, and to a lesser extent in the

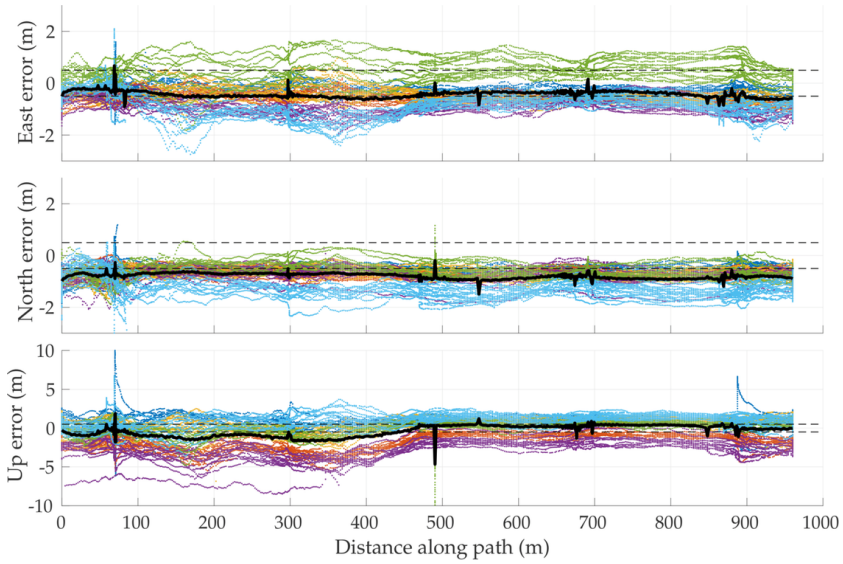


Figure 5.8: Errors in enhanced code-phase position estimates with respect to ground truth in the east, north, and up directions. Different colors distinguish data from six different campaigns. The dashed reference lines are drawn at $\pm$ 50 cm. The solid black lines show the mean positioning error over the six campaigns. The error standard deviation is nearly constant along the path in the horizontal plane at $\sim$0.6 m in the east and $\approx$0.4 m in the north direction. In the up direction, the standard deviation is $\sim$2.1 m for the first 400 m along the path, and $\approx$1.3 m for the rest.

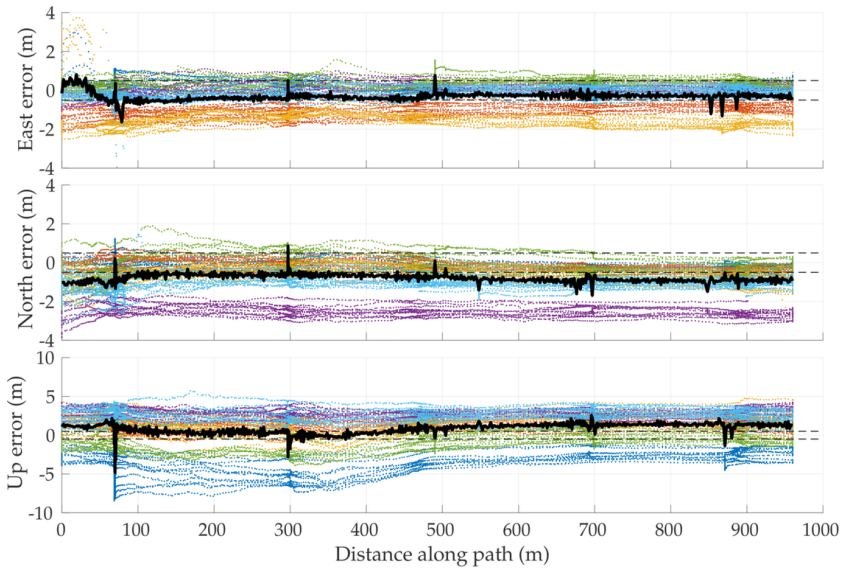


Figure 5.9: Errors in ublox M8T position estimates with respect to ground truth in the east, north, and up directions. Different colors distinguish data from six different campaigns. Dashed reference lines are drawn at ± 50 cm. The solid black lines show the mean positioning error over the six campaigns. The error standard deviation in the east is ∼1.5 m over the first 100 m along the path and ∼0.7 m over the rest; ∼0.9 m in the north; and ∼2.7 m over the first 400 m and ∼2 m over the rest in the up direction.

east, may only be an artifact of the small sample size: ionospheric modeling errors were not yet averaged down to nearly zero in the east and ∼30 cm in the north, as one would expect from the WAAS ionospheric model (see Table 5.1).

Given that the asymptotic properties of ionospheric modeling errors are better understood than those of multipath errors, it is instructive to eliminate, insofar as possible, all ionospheric modeling errors from the along-track error histories. To this end, a differential code phase GNSS technique is applied whereby the navigation solution was based on double-difference pseudorange measurements using data from a nearby reference station at a precisely known location. Such double differencing over a short 1-km baseline eliminates virtually all ionospheric and tropospheric errors, but does nothing to reduce vehicle-side multipath. Thus, one can empirically examine multipath effects in isolation from ionospheric effects.

Fig. 5.10 shows the results of this study based on all six data capture campaigns. Note that biases for all components are much smaller. It appears that for the test route chosen, non-zero-mean horizontal errors in the enhanced code phase positions are almost entirely driven by ionospheric modeling errors, and not by persistent effects of multipath due to the permanent structures along the test route. This is broadly consistent with the analyses presented earlier in this chapter on position-domain biases due to ionospheric and multipath errors. However, it does appear that a bias due to multipath remained in the vertical direction over the first 400 m, even when ionospheric errors were removed. Apparently, the arrangement of buildings over this segment

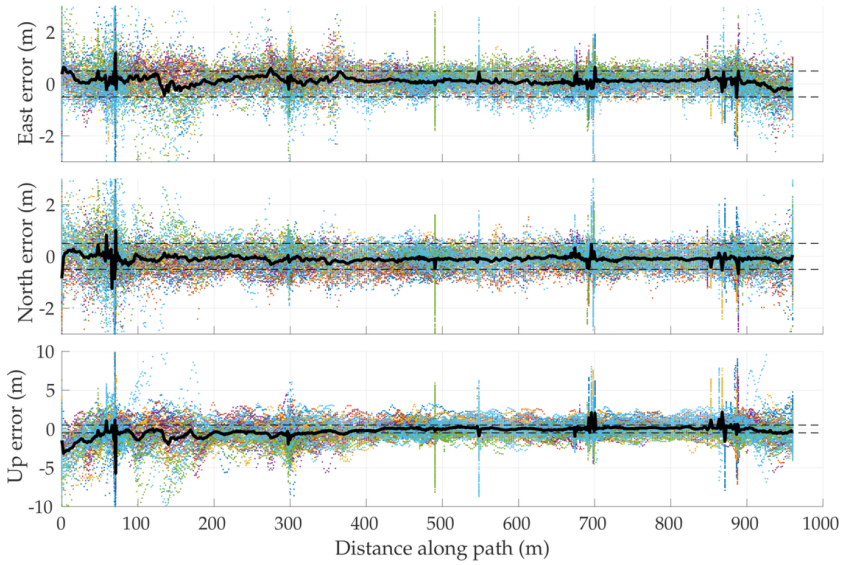


Figure 5.10: Errors in double-differenced pseudorange-based position estimates with respect to ground truth in the east, north, and up directions. Different colors distinguish data from six different campaigns. Dashed reference lines are drawn at ±50 cm. The solid black lines show the mean positioning error over the six campaigns. The error standard deviation in the east and north directions is ∼0.9 m over the first 200 m along the path and ∼0.4 m over the rest. In the up direction, the standard deviation is ∼1.9 m over the first 400 m and ∼1 m over the rest.

caused non-line-of-sight effects that did not average away. Mercifully, horizontal errors, which appear to be close to zero-mean over the six campaigns, matter most for high-accuracy digital mapping, since obstacle avoidance and vehicle coordination are largely 2-D problems, and since multiple vehicles can straightforwardly agree on a particular feature's relative vertical position from an inferred road surface.

Based on Fig. 5.10, one can conclude that multi-session averaging with a sufficiently accurate ionospheric model, such as the Fast PPP model, yields

sub-50-cm global referencing accuracy for digital maps in the horizontal plane with code-phase-based GNSS, even in the presence of persistent multipath.

5.6 Conclusion

The accuracy limits of collaborative global referencing of digital maps with standard GNSS were explored through modeling and simulation. The asymptotic average of position errors due to thermal noise, satellite orbit and clock errors, and tropospheric modeling errors were assumed to be negligible. From Sec. 5.4.3.3's analysis of asymptotic ionospheric errors, and from Sec. 5.4.3.5's multipath simulation study, one can draw the following conclusion: so long as the asymptotic horizontal position errors of the ionosphere corrections are below 5 cm, as is true for the Fast-PPP model, and assuming statistical independence of ionospheric and multipath errors, it appears feasible to achieve 50-cm horizontal positioning accuracy at approximately 95% by averaging over 100 mapping sessions. Empirical results from a field experiment involving 75 laps of a 1 km loop support this conclusion.

Chapter 6

Globally-referenced Electro-Optical SLAM

6.1 Abstract

This chapter presents a globally-referenced electro-optical simultaneous localization and mapping pipeline, called GEOSLAM, designed for crowd-sourced mapping and localization. This pipeline serves as a demonstration of a system that achieves the asymptotic accuracy limit of collaborative digital mapping described in Chap. 5. GEOSLAM achieves this accuracy by (i) incorporating standard GNSS position estimates in the visual simultaneous localization and mapping (SLAM) framework, (ii) merging digital maps from multiple mapping sessions, and (iii) jointly optimizing structure and motion with respect to time-separated GNSS measurements. Field experiments are conducted in a moderately-urban environment to show that after 8 sessions of joint optimization, GEOSLAM generates a map of visual features with 50 cm global accuracy.

6.2 Introduction

Mapping of the static driving environment is key to robust positioning and navigation systems, including the radar-based positioning engine described

in Chap. 4. These so-called *high-definition (HD) maps* for AGVs have many other applications beyond localization: responding to traffic signs and signals, for example, is greatly simplified if the vehicle has prior knowledge of where such signs are located. In short, a map of the surrounding environment enables an AGV to *expect the expected*.

Generating and maintaining these HD maps, however, is a major challenge. Most AGV manufacturers, such as Waymo and General Motors, deploy specialized fleets of mapping vehicles. Generating a map of the environment requires precise knowledge of the vehicle pose (position and orientation) that must be obtained with either an expensive tactical-grade INS or with a high-resolution lidar in a SLAM framework, or a combination thereof. Furthermore, the map must be updated whenever the static environment changes, e.g., due to construction. It is time-consuming and impractical to maintain HD maps of entire continents.

A key enabler for large-scale up-to-date maps will be enlisting the help of the very consumer vehicles that need the map to build and update it. Consumer vehicles are typically equipped with low-cost consumer-grade sensor suites. As mentioned in Chap. 5, it is likely that different car manufacturers would not share a common HD map, and that data sharing for increased situational-awareness would require all such maps to be consistent in a global frame of reference. Chap. 5 further claimed that such maps can be crowd-sourced from the consumer vehicles, provided that the map generation system can jointly optimize the crowd-sourced data over time-separated measurements

from standard code-phase-based GNSS receivers.

The goal of this chapter is to describe and demonstrate one such pipeline named GEOSLAM (globally-referenced electro-optical simultaneous localization and mapping) that is capable of globally-referenced collaborative multi-session digital mapping. The pipeline combines visual measurements from a stereo visible-light camera system with position measurements from GNSS signals. The basic intuition is that if one or more camera-equipped vehicles navigating through a digital map over time make multiple time-separated GNSS measurements of the same location, then the GNSS errors can be averaged across all sessions with appropriate weighting.

As such, visual SLAM is a mature field of research. The contributions of this chapter deal with the integration of GNSS measurements in visual SLAM, and with other challenges that arise specifically in the case of multi-session or crowd-sourced mapping and localization. This chapter details the techniques GEOSLAM invokes to smoothly transition between unmapped and previously-mapped regions, consistently fusing current and prior maps without the need for a six degrees-of-freedom (6-DoF) pose from an INS. GEOSLAM enables multi-agent collaborative mapping by storing and rendering its map in a global frame of reference, such as the World Geodetic System 1984 or the International Terrestrial Reference Frame. Multi-session operation of GEOSLAM is demonstrated using camera and GNSS data collected in a moderate urban environment, and the accuracy of global localization with the multi-session GEOSLAM map is assessed with respect to CDGNSS-based ground truth.

6.3 Related Work

Sensor fusion of visible-light cameras and GNSS has been extensively studied [10,29,39,49,76,84,85,118,141,142,147,162]. Some of these works [10, 39,162] have proposed visual odometry as a replacement for, or an augmentation of, the traditional GNSS/INS architecture. Visual data from cameras are exploited to perform dead reckoning in a visual odometry pipeline, wherein an important distinction from the current chapter is that the 3D map points do not persist after a window of time has elapsed—that is, no map of feature points is maintained. Clearly, such an approach does not allow improvement of the 3D map point positions over multiple mapping sessions.

In [147], the relative change in position between two image frames is first estimated based on time-differenced GNSS carrier phase measurements. The metric-accurate GNSS-derived change in position is exploited to initialize the otherwise unobservable scale in a monocular visual SLAM system. However, GNSS measurements are not incorporated any further once the absolute scale has been initialized. Unlike the current chapter, the visual SLAM map is rendered in the arbitrary SLAM coordinate frame since only the relative change in position, and not the absolute position, was estimated based on GNSS measurements.

The vision-GNSS fusion in the current chapter is closely aligned with the bi-objective bundle adjustment (BA) optimization techniques previously reported in [29, 76, 84, 85, 141, 142]. In [29, 76, 141, 142], the traditional visual SLAM reprojection cost function is jointly minimized along with a GNSS

position error term. The methods proposed in [84, 85] are also similar, but guarantee that the visual reprojection cost after incorporation of the GNSS term is not significantly greater than the visual-only case. However, none of these works showed significant empirical evidence of their efficacy on real-world vehicle data sets. Furthermore, collaborative mapping or multi-session improvement of the map was not discussed.

Collaborative multi-agent mapping, without GNSS aiding, has also been extensively discussed in the literature [6, 44, 55, 69, 121, 131, 177]. Some of these proposed solutions require significant overlap in the field-of-view of the agents, or require that the relative pose transformation between the agents be known a priori [121, 177]. Other solutions, such as in [6, 44, 55, 69], enable collaboration by performing data association between non-concurrent mapping sessions where the relative pose transformation between the agents is unknown. The multi-session strategy employed in this chapter is similar, but with an important distinction: none of the previous works on collaborative mapping have incorporated GNSS measurements in the map-building process. Without global referencing, the problem of data association between non-concurrent sessions becomes intractable. With no estimate of the pose for the mapping platform in relation to the existing map, data association must be attempted against the entire map. It is easily observed that such data association will become infeasible when scaled to city- or country-wide maps. The current chapter proposes rendering and storage of digital maps in a global coordinate frame, such that a new mapping session can readily estimate its approximate

pose in relation to the prior map, and perform data association on a small segment of the prior map that is expected to be in view of the vision system.

The work presented in [45] is perhaps the most closely related to the current chapter. In [45], a particular stretch of roadway is mapped 25 times with a low-cost sensor setup. However, [45] assumes, without detail, the availability of a lane-level accurate low-cost positioning module that provides the full 6-DoF pose for the mapping platform. This greatly simplifies the ensuing data association and mapping pipeline. No mention is made of the general setting of the roadway being mapped (open-sky highway, urban canyon, etc.), and while the accuracy of the mapped traffic signs is adequately reported, localization within the map is not discussed, presumably since lane-level accurate positioning is already available. Meanwhile, the current chapter only assumes the availability of a meter-level accurate code-phase-based GNSS receiver that provides 3D position estimates. Global localization accuracy of a vehicle operating within the multi-session map is presented as the key performance indicator.

6.4 Visual SLAM

The visual SLAM component of GEOSLAM is similar to existing high-performance SLAM pipelines developed in the robotics community [50,81,103]. Visual SLAM algorithms may be categorized as either sparse or dense. Sparse visual SLAM algorithms [81, 103] create a map of distinctive features such as corners or edges in the scene, while dense SLAM algorithms [50] map the

depth for each pixel in the captured frames. The point cloud generated by sparse SLAM algorithms is sufficient for the purpose of localization. Dense reconstruction is appealing to the human eye, but does not provide any tangible benefit to localization, while consuming much more computational resources. As a result, GEOSLAM implements sparse feature-point-based SLAM.

In [151] it was shown that for the visual SLAM problem, structure-from-motion BA (batch non-linear optimization) outperforms filtering techniques such as the extended Kalman filter, yielding higher accuracy per unit of computing time. It was also noted that having a high number of features points per image frame provides better accuracy than having a large number of frames with fewer feature points per frame. Thus, in typical practice, only a select subset of frames among those captured is retained for processing; frames in this subset are called keyframes. Most recent state-of-the-art visual SLAM algorithms use a keyframe-based BA approach instead of sequential filtering. Likewise, GEOSLAM performs BA-based non-linear optimization to refine both structure and motion.

Fig. 6.1 shows a block diagram representation of the system architecture proposed in this chapter. The yellow-colored blocks in this figure are components of the GEOSLAM pipeline, detailed next.

By way of notation, let

$$\boldsymbol{z}_{nm}^{lr} \triangleq \left(u_{nm}^{l}, v_{nm}^{l}, u_{nm}^{r}\right)$$

denote the image plane location of the stereo-matched feature matched to the

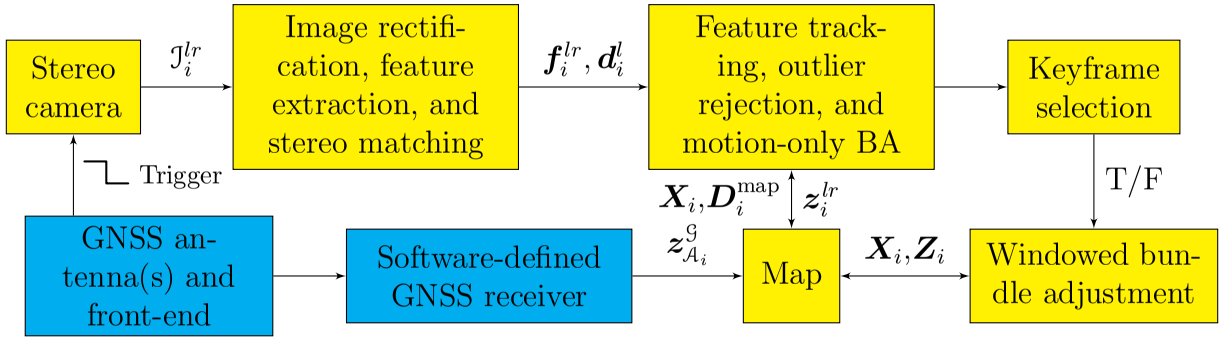


Figure 6.1: Globally-referenced electro-optical simultaneous localization and mapping (GEOSLAM) block diagram. BA: bundle adjustment.

mth map point, p_m, in the nth stereo keyframe, K_n. The horizontal and vertical coordinates are denoted u and v, respectively, while the superscripts l and r denote the left and right camera frames, respectively. Note that the feature location is specified by only three coordinates. The vertical feature coordinate in the right camera frame, v_{nm}^r, is omitted because for an undistorted and rectified camera model it must hold that $v_{nm}^l = v_{nm}^r$, making one of the coordinates redundant. If p_m is not matched to any feature in K_n, then let $\boldsymbol{z}_{nm}^{lr} = \varnothing$. Furthermore, let

$$\mathcal{M}_n = \left\{ m \ : \ \boldsymbol{z}_{nm}^{lr} \neq \varnothing \right\}$$

denote the set of indices of all map points matched to some feature in K_n. In the visual SLAM literature, the covisibility window of keyframe K_i is defined as the set of keyframes that share at least T map points with K_i. Mathematically, the covisibility window of keyframe K_i is the set of keyframes with indices

$$\text{cov}(i) \triangleq \{ n \ : \ |\mathcal{M}_n \cap \mathcal{M}_i| > T \},$$

where $|\mathcal{A}|$ denotes the cardinality of the set $\mathcal{A}$. The covisibility window determines the keyframes to be optimized in a windowed BA. The visibility of common points is regarded as a proxy for correlation between the structure-from-motion states. However, in a sensor fusion architecture, the states for other sensors (e.g., GNSS) may be spatially correlated beyond the covisibility window. Furthermore, other sensors may experience outages that extend beyond the covisibility window. In such a scenario, it would be desirable to

optimize over a batch of keyframes that span the availability gap. Accordingly, GEOSLAM extends the concept of covisibility to N levels as

$$\mathrm{cov}(i,N) \triangleq \begin{cases} \left\{n \ : \ |\mathcal{M}_n \cap \left(\cup_{k\in\mathrm{cov}(i,N-1)}\mathcal{M}_k\right)| > T\right\} & N>1, \\ \{n \ : \ |\mathcal{M}_n \cap \mathcal{M}_i| > T\} & N=1. \end{cases}$$

When processing K_i, GEOSLAM's objective is to estimate the map point 3D locations $\boldsymbol{x}^{\mathcal{S}}_{p_m} \in \mathbb{R}^3$ and keyframe poses $\left(\boldsymbol{x}^{\mathcal{S}}_{\mathcal{C}_n}, \boldsymbol{\theta}^{\mathcal{S}}_{\mathcal{C}_n}\right) \in (\mathbb{R}^3, \mathbb{R}^3)$ in the N-level covisibility window for the ith keyframe, where $\mathcal{S}$ stands for the local SLAM frame, $\mathcal{C}_n$ is the left camera coordinate frame associated with K_n, and $\boldsymbol{\theta}^{\mathcal{S}}_{\mathcal{C}_n}$ is the angle-axis representation of the keyframe orientation. The state vector to be estimated is represented as

$$\boldsymbol{X}_i \triangleq \left[\left\{\left(\boldsymbol{x}^{\mathcal{S}}_{\mathcal{C}_n}, \boldsymbol{\theta}^{\mathcal{S}}_{\mathcal{C}_n}\right) \ : \ n \in \mathrm{cov}(i,N)\right\}, \left\{\boldsymbol{x}^{\mathcal{S}}_{p_m} \ : \ m \in \cup_{k\in\mathrm{cov}(i,N)}\mathcal{M}_k\right\}\right]^T, \quad (6.1)$$

where the two sets on the right-hand side are arranged as a concatenation of row vectors so that $\boldsymbol{X}_i$ becomes a column vector.

When triggered by the GNSS front end, the camera setup captures a pair, denoted $\mathcal{I}^{lr}_i$, of concurrent images from the left and right cameras, where the subscript denotes that the current pair is a candidate to be the ith stereo keyframe K_i. The intrinsic and extrinsic parameters of the stereo camera setup are assumed to have been calibrated a priori. The stereo image pair is then undistorted and rectified according to the given calibration, and SIFT features are detected and computed separately for each image [89]. SIFT feature matching is performed between the left and right image with the additional constraint that matching features must have approximately the same vertical

coordinate to within a few pixels. The set of stereo feature measurements for $\mathcal{I}_i^{lr}$, $\boldsymbol{f}_i^{lr}$, and the set of feature descriptors as computed in the left image, $\boldsymbol{d}_i^l$, are passed on to the tracking module.

The tracking module has access to the 3D map point positions within $\boldsymbol{X}_i$ and to the set of SIFT descriptors, $\boldsymbol{D}_i^{\text{map}}$, corresponding to the map points expected to be seen in the candidate keyframe $\mathcal{I}_i^{lr}$. The tracker performs directed matching of the features between the stereo image and the map. First, a quick feature matching is performed using the Fast Approximate Nearest Neighbor Search Library (FLANN) [100]. With sufficient matches, an initial approximation of the current camera pose is obtained using the five-point algorithm wrapped in RANSAC iterations. With this approximate pose, an iteration of exhaustive nearest neighbor search is performed for each map point potentially in view of the camera, but only within a small window of its projected position on the image plane. Subsequently, RANSAC iterations are performed on the full set of feature matches to remove any remaining outliers, and a motion-only BA is performed wherein the current camera pose is optimized based on the feature matches to a fixed set of 3D map points.

After tracking the stereo image pair as described above, GEOSLAM decides whether or not the candidate keyframe $\mathcal{I}_i^{lr}$ must be selected as a keyframe. This decision is made based on the number of map points that were matched to the image features, and the distance traveled by the platform since the last keyframe was chosen. New keyframes are not spawned if the platform is nearly stationary. If the platform is in motion, and the number of feature matches to

the map drops below a threshold, then the candidate keyframe $\mathcal{I}_i^{lr}$ is chosen as keyframe K_i, windowed BA is performed over N levels of covisibility, and the unmatched stereo features in K_i are spawned as new map points. Additionally, if $\mathcal{I}_i^{lr}$ is selected to be K_i, then the set of measurements from the features matched to the map, denoted $\boldsymbol{z}_i^{lr} \triangleq \{\boldsymbol{z}_{im}^{lr} \ : \ m \in \mathcal{M}_i\}$, along with their SIFT descriptors, are passed on to the map module for storage, future feature matching, and processing in the windowed BA routine.

In a visual-only SLAM system, the state vector $\boldsymbol{X}_i$ is optimized with respect to the measurement vector $\boldsymbol{Z}_i$, defined as

$$\boldsymbol{Z}_i \triangleq \left[\{\boldsymbol{z}_n^{lr} \ : \ n \in \text{cov}(i, N)\}\right]^T.$$

The windowed BA routine in GEOSLAM minimizes the 3D-to-2D reprojection error. The error term $\boldsymbol{e}_{nm}$ for observation of map point p_m in the stereo keyframe K_n is given as

$$\boldsymbol{e}_{nm} = \boldsymbol{z}_{nm}^{lr} - \Pi(\boldsymbol{x}_{p_m}^{\mathcal{S}}, \boldsymbol{x}_{\mathcal{C}_n}^{\mathcal{S}}, \boldsymbol{\theta}_{\mathcal{C}_n}^{\mathcal{S}}),$$

where Π is the projection function for an undistorted and rectified stereo camera

$$\Pi(\boldsymbol{x}_{p_m}^{\mathcal{S}}, \boldsymbol{x}_{\mathcal{C}_n}^{\mathcal{S}}, \boldsymbol{\theta}_{\mathcal{C}_n}^{\mathcal{S}}) = \begin{bmatrix} f\frac{x_{nm}}{z_{nm}} + c_u \\ f\frac{y_{nm}}{z_{nm}} + c_v \\ f\frac{x_{nm}}{z_{nm}} + c_u - bf \end{bmatrix},$$
$$[x_{nm}, y_{nm}, z_{nm}]^T = R(\boldsymbol{\theta}_{\mathcal{C}_n}^{\mathcal{S}})^T(\boldsymbol{x}_{p_m}^{\mathcal{S}} - \boldsymbol{x}_{\mathcal{C}_n}^{\mathcal{S}}),$$

in which $R(\cdot)$ denotes the rotation matrix corresponding to the argument angle-axis vector, and f, (c_u, c_v), and b are the focal length, the principal

point, and the baseline distance between the left and right cameras of the rectified stereo camera model, respectively. The cost function to be minimized for visual SLAM is given as

$$C_i = \sum_{n \in \text{cov}(i,N)} \sum_{m \in \mathcal{M}_n} \rho\big(\boldsymbol{e}_{nm}^T \Omega_{nm}^{-1} \boldsymbol{e}_{nm}\big),$$

where ρ may be the standard least squares cost function $\rho(\cdot) = (\cdot)$ or a more robust cost function such as the Huber or Tukey cost functions, and where $\Omega_{nm} = \sigma_{nm}^2 I_{3\times 3}$ is the covariance of the feature measurements.

GEOSLAM performs BA minimization via Google's `ceres-solver`. The automatic differentiation feature of `ceres-solver` is used to compute the Jacobian for the measurement model.

An important feature of the GEOSLAM visual pipeline is the ability to merge maps from multiple mapping sessions. This is embodied in an algorithm similar to the loop closure technique from the visual SLAM literature [102]. Map merging is described in detail in Sec. 6.6.2.

6.5 GNSS Aiding

Conventional visual SLAM algorithms are known to drift from the true platform trajectory as a function of the distance traveled by the platform. Furthermore, the map of the structure is created in the arbitrary $\mathcal{S}$ frame. Such a map cannot be intelligibly shared with another mapping agent having a different $\mathcal{S}$ frame. Meanwhile, GNSS position estimates are obtained in the global $\mathcal{G}$ frame and do not exhibit any distance-dependent drift. Accordingly,

GEOSLAM ingests standard GNSS position estimates from a software-defined GNSS receiver, called GRID/pprx [64,87], in a tightly-coupled architecture to create a globally-referenced map, enable cooperative multi-session mapping, and constrain the drift of visual SLAM. Since the stereo camera setup is triggered by the same clock that drives digitization of the GNSS samples (see Fig. 6.1), it is possible to produce GNSS measurements synchronized with the camera image epochs. This section details the various coordinate frames in GNSS-aided visual SLAM, the updated BA cost function, and an initialization routine required to enable GNSS aiding in SLAM.

6.5.1 Coordinate Frames

The GNSS-aided visual SLAM system has three coordinate frames of interest: the K_i camera frame $\mathcal{C}_i$, the local SLAM frame $\mathcal{S}$, and the global frame $\mathcal{G}$. The SLAM frame $\mathcal{S}$ adopts the position and orientation of the first keyframe prior to optimization as its origin and orientation. Thus, $\mathcal{S}$ is fixed relative to $\mathcal{G}$, but each $\mathcal{C}_i$ changes relative to $\mathcal{G}$ as the platform moves.

Note that the structure-from-motion states in Equation (6.1) are represented in the $\mathcal{S}$ frame, whereas the K_nth keyframe's corresponding GNSS measurement, denoted $\boldsymbol{z}^{\mathcal{G}}_{\mathcal{A}_n} \in \mathbb{R}^3$, is natively represented in the $\mathcal{G}$ frame. The latter is transformed to the $\mathcal{S}$ frame through an unknown but fixed rotation, $R^{\mathcal{S}}_{\mathcal{G}} \in SO(3)$, and translation, $\boldsymbol{t}^{\mathcal{S}}_{\mathcal{G}} \in \mathbb{R}^3$. This transformation is estimated at initialization as explained in Sec. 6.5.2. After initialization, $\boldsymbol{z}^{\mathcal{G}}_{\mathcal{A}_n}$ is rendered in

$\mathcal{S}$ as

$$\boldsymbol{z}^{\mathcal{S}}_{\mathcal{A}_n} = R^{\mathcal{S}}_{\mathcal{G}} \boldsymbol{z}^{\mathcal{G}}_{\mathcal{A}_n} + \boldsymbol{t}^{\mathcal{S}}_{\mathcal{G}}.$$

GEOSLAM estimates the 6-DoF pose of the left camera, but the GNSS antenna phase center is not co-located with the camera center; rather, it is offset from the camera center by a fixed vector (same for all n) in $\mathcal{C}_n$ denoted, $\boldsymbol{t}^{\mathcal{C}_n}_{\mathcal{A}_n} \in \mathbb{R}^3$. Thus, the error term associated with the GNSS position estimate for K_n is given as

$$\boldsymbol{e}_{\mathcal{A}_n} = \boldsymbol{z}^{\mathcal{S}}_{\mathcal{A}_n} - \big(\boldsymbol{x}^{\mathcal{S}}_{\mathcal{C}_n} + R\big(\boldsymbol{\theta}^{\mathcal{S}}_{\mathcal{C}_n}\big)\boldsymbol{t}^{\mathcal{C}_n}_{\mathcal{A}_n}\big).$$

Under the assumption of temporally-uncorrelated GNSS errors, the updated BA cost function to be minimized is

$$C_i = \sum_{n \in \mathrm{cov}(i,N)} \left[\sum_{m \in \mathcal{M}_n} \rho\big(\boldsymbol{e}^T_{nm} \Omega^{-1}_{nm} \boldsymbol{e}_{nm}\big) + \boldsymbol{e}^T_{\mathcal{A}_n} \Gamma^{-1}_n \boldsymbol{e}_{\mathcal{A}_n} \right],$$

where $\Gamma_n = R^{\mathcal{S}}_{\mathcal{G}} \Gamma'_n \big(R^{\mathcal{S}}_{\mathcal{G}}\big)^T$ and Γ'_n is the covariance matrix of the GNSS position estimate associated with K_n, expressed in $\mathcal{G}$.

6.5.2 Initialization in GNSS-Aided SLAM

When initializing, GEOSLAM performs visual-only SLAM for the first N_i keyframes in the $\mathcal{S}$ frame, and stores the GNSS position measurements of the antenna provided in the $\mathcal{G}$ frame. Subsequently, GEOSLAM finds the least-squares Euclidean transformation to obtain the optimal rotation matrix $R^{\mathcal{S}}_{\mathcal{G}}$ and translation vector $\boldsymbol{t}^{\mathcal{S}}_{\mathcal{G}}$ between the two coordinate systems from the set of vector observations. Note that a full similarity transformation is not

required since the known stereo baseline renders the $\mathcal{S}$ frame with correct scaling. The estimated Euclidean transformation minimizes the squared difference between the transformed GNSS measurements in $\mathcal{S}$ and the visual SLAM predicted trajectory of the GNSS antenna, also in $\mathcal{S}$. The specific method used to estimate the transformation is based on SVD decomposition as discussed in [148].

$$\left(R^{\mathcal{S}}_{\mathcal{G}}, \boldsymbol{t}^{\mathcal{S}}_{\mathcal{G}}\right) = \underset{R\in SO(3), \boldsymbol{t}\in\mathbb{R}^3}{\operatorname{argmin}} \sum_{n=1}^{N_i} \left\| \left(R\boldsymbol{z}^{\mathcal{G}}_{\mathcal{A}_n} + \boldsymbol{t}\right) - \left(\boldsymbol{x}^{\mathcal{S}}_{\mathcal{C}_n} + R\left(\boldsymbol{\theta}^{\mathcal{S}}_{\mathcal{C}_n}\right)\boldsymbol{t}^{\mathcal{C}_n}_{\mathcal{A}_n}\right) \right\|^2 \tag{6.2}$$

It must be noted that this transformation need only be approximately correct such that the GNSS estimates, used as measurements, will not diverge with respect to the visually-derived trajectory. Because the jointly estimated trajectory in $\mathcal{S}$ gets transformed back to $\mathcal{G}$ using the same approximate transformation, any errors in the transformation are cancelled.

6.6 Multi-Session Mapping

Refinement of the visual feature map over multiple sessions with time-separated GNSS measurements is central to the idea of approaching the accuracy limit of mapping with standard GNSS. Consider a vehicle revisiting an area mapped previously in one or more sessions. When GEOSLAM matches greater than T features in the current keyframe to the features already present in the prior map, the keyframes from the previous sessions in that section of the map are included in the covisibility window of the current keyframe. After such a merge is detected and verified, a BA may be performed on the covisible

keyframes from multiple sessions to average time-separated standard GNSS errors. It is important to note that multi-session mapping can only be realized when sufficient feature matches are found between multiple sessions. This is not a straightforward task, as evidenced by recent efforts on lifelong feature mapping efforts [99]. This issue is further discussed in Sec. 6.7. In the current section, multi-session map database management and map merging are discussed.

6.6.1 Map Database

Storage and reuse of maps is a pre-requisite for multi-session mapping. For a given session, the SLAM map is created in the $\mathcal{S}$ frame. However, such a map is not readily usable in successive mapping sessions since the $\mathcal{S}$ frame is distinct for each session. Fortunately, the integration of visual SLAM with GNSS enables transformation of the SLAM map in to the $\mathcal{G}$ frame.

At the end of the pth mapping session in the local frame $\mathcal{S}_p$, GEOSLAM stores the data to a map database after applying the $\left(R_{\mathcal{G}}^{\mathcal{S}_p}, \boldsymbol{t}_{\mathcal{G}}^{\mathcal{S}_p}\right)$ transformation, as estimated during initialization for the pth session, to all map point positions, all keyframe poses, and all GNSS measurements associated with each keyframe:

$$
\begin{aligned}
\boldsymbol{x}_{\mathcal{C}_n}^{\mathcal{G}} &= \left(R_{\mathcal{G}}^{\mathcal{S}_p}\right)^T \left(\boldsymbol{x}_{\mathcal{C}_n}^{\mathcal{S}_p} - \boldsymbol{t}_{\mathcal{G}}^{\mathcal{S}_p}\right); \quad R\left(\boldsymbol{\theta}_{\mathcal{C}_n}^{\mathcal{G}}\right) = \left(R_{\mathcal{G}}^{\mathcal{S}_p}\right)^T R\left(\boldsymbol{\theta}_{\mathcal{C}_n}^{\mathcal{S}_p}\right), \\
\boldsymbol{x}_{p_m}^{\mathcal{G}} &= \left(R_{\mathcal{G}}^{\mathcal{S}_p}\right)^T \left(\boldsymbol{x}_{p_m}^{\mathcal{S}_p} - \boldsymbol{t}_{\mathcal{G}}^{\mathcal{S}_p}\right), \\
\boldsymbol{z}_{\mathcal{C}_n}^{\mathcal{G}} &= \left(R_{\mathcal{G}}^{\mathcal{S}_p}\right)^T \left(\boldsymbol{z}_{\mathcal{C}_n}^{\mathcal{S}_p} - \boldsymbol{t}_{\mathcal{G}}^{\mathcal{S}_p}\right).
\end{aligned}
$$

At the beginning of the $(p+1)$th session, GEOSLAM again estimates

the $\left(R_{\mathcal{G}}^{\mathcal{S}_{p+1}}, \boldsymbol{t}_{\mathcal{G}}^{\mathcal{S}_{p+1}}\right)$ transformation during initialization. The map database from previous session(s) is then loaded after applying the transformation for the $p+1$ session, such that the prior map points, keyframes, and measurements are rendered in the $\mathcal{S}_{p+1}$ frame:

$$\begin{aligned}
\boldsymbol{x}_{\mathcal{C}_n}^{\mathcal{S}_{p+1}} &= R_{\mathcal{G}}^{\mathcal{S}_{p+1}} \boldsymbol{x}_{\mathcal{C}_n}^{\mathcal{G}} + \boldsymbol{t}_{\mathcal{G}}^{\mathcal{S}_{p+1}}; \quad R\left(\boldsymbol{\theta}_{\mathcal{C}_n}^{\mathcal{S}_{p+1}}\right) = \left(R_{\mathcal{G}}^{\mathcal{S}_{p+1}}\right) R(\boldsymbol{\theta}_{\mathcal{C}_n}^{\mathcal{G}}), \\
\boldsymbol{x}_{p_m}^{\mathcal{S}_{p+1}} &= R_{\mathcal{G}}^{\mathcal{S}_{p+1}} \boldsymbol{x}_{p_m}^{\mathcal{G}} + \boldsymbol{t}_{\mathcal{G}}^{\mathcal{S}_{p+1}}, \\
\boldsymbol{z}_{\mathcal{C}_n}^{\mathcal{S}_{p+1}} &= R_{\mathcal{G}}^{\mathcal{S}_{p+1}} \boldsymbol{z}_{\mathcal{C}_n}^{\mathcal{G}} + \boldsymbol{t}_{\mathcal{G}}^{\mathcal{S}_{p+1}}.
\end{aligned}$$

After loading the prior map, the standard GEOSLAM pipeline is executed for each stereo image pair in the $(p+1)$th session. In addition, GEOSLAM attempts to detect if the vehicle is currently passing through a previously-mapped region. If so, a map merge is declared and the current and prior keyframes are jointly optimized, as detailed in Sec. 6.6.2. Finally, at the end of the mapping session, the combined map is stored back in the database as described before.

6.6.2 Map Merging

As mentioned before, the matching of feature points across multiple sessions is central to the idea of averaging standard GNSS errors. Once sufficiently many features are matched between the current stereo keyframes and prior map points, GEOSLAM declares a map merging event. This is akin to the well-known problem of detecting loop closure in the visual SLAM literature [102]. This section details GEOSLAM's map merging and loop closing

routine. Hereafter, the terms map merging and loop closure are used interchangeably since GEOSLAM treats them identically.

First, note that when detecting a map merge event, feature matching must be attempted against map points that have not been matched in the most recent keyframes. Thus, after processing the ith keyframe, a possible merge is checked for against the set of map points $\{m \ : \ m \notin \cup_{n \in \mathrm{cov}(i,N)} \mathcal{M}_n\}$. If this bag-of-words-style feature matching succeeds, then RANSAC iterations are performed to determine whether the matches are geometrically consistent, as well as to robustly estimate the camera pose $\left(\check{\boldsymbol{x}}^{\mathcal{S}}_{\mathcal{C}_i}, \check{\boldsymbol{\theta}}^{\mathcal{S}}_{\mathcal{C}_i}\right)$ implied by the merge event. If enough inliers are found, the map merge routine is executed.

The map merging process is depicted visually in Figs. 6.2–6.4. A typical merge situation is shown in Fig. 6.2, where the platform pose at the ith keyframe is inconsistent with the prior map at the merge location. To avoid such discontinuity in the ensuing joint BA, as an initial guess GEOSLAM enforces that the ith keyframe pose be consistent with the pose implied by the visual merge matches $\left(\check{\boldsymbol{x}}^{\mathcal{S}}_{\mathcal{C}_i}, \check{\boldsymbol{\theta}}^{\mathcal{S}}_{\mathcal{C}_i}\right)$, and that the keyframes and map points from the prior session(s) be unchanged. To this end, a pose-graph optimization [150] is performed over a large N_m-level covisibility window for K_i, where the relative translations and rotations between covisible keyframes, as estimated in the current session, are provided as delta-pose measurements, while the 6-DoF poses for the terminal nodes in the covisibility window, as well as for K_i, are held constant. In particular, let $\mathcal{K}_0$ denote the set of terminal keyframes in the covisibility graph $\mathrm{cov}(i, N_m)$. Furthermore, define the delta-pose pseudo-

measurements

$$\delta \boldsymbol{x}^{\mathcal{S}}_{nk} \triangleq \hat{\boldsymbol{x}}^{\mathcal{S}}_{\mathcal{C}_n} - \hat{\boldsymbol{x}}^{\mathcal{S}}_{\mathcal{C}_k},$$
$$\delta \boldsymbol{\theta}^{\mathcal{S}}_{nk} \triangleq \theta\Big(R\Big(\hat{\boldsymbol{\theta}}^{\mathcal{S}}_{\mathcal{C}_n}\Big) R\Big(\hat{\boldsymbol{\theta}}^{\mathcal{S}}_{\mathcal{C}_k}\Big)^T \Big),$$

where the superscript $(\hat{\cdot})$ denotes GEOSLAM's estimate of the state before the merge event, and $\theta(\cdot)$ denotes the angle-axis representation of the input rotation matrix. The pose-graph optimization minimizes the following cost function with respect to $\big(\boldsymbol{x}^{\mathcal{S}}_{\mathcal{C}_n}, \boldsymbol{\theta}^{\mathcal{S}}_{\mathcal{C}_n}\big) \forall n \in \text{cov}(i, N_m)$:

$$C = \sum_{n \in \text{cov}(i,N_m)} \sum_{k \in \text{cov}(n,1)} \Big[\Big\| \big(\boldsymbol{x}^{\mathcal{S}}_{\mathcal{C}_n} - \boldsymbol{x}^{\mathcal{S}}_{\mathcal{C}_k}\big) - \delta \boldsymbol{x}^{\mathcal{S}}_{nk} \Big\|^2_{P^{-1}_{\delta \boldsymbol{x}}} + \Big\| \theta\Big(R\big(\boldsymbol{\theta}^{\mathcal{S}}_{\mathcal{C}_n}\big) R\big(\boldsymbol{\theta}^{\mathcal{S}}_{\mathcal{C}_{nk}}\big)^T \Big) - \delta \boldsymbol{\theta}^{\mathcal{S}}_{n-1} \Big\|^2_{P^{-1}_{\delta \boldsymbol{\theta}}} \Big],$$

where $\|\boldsymbol{q}\|^2_P = \boldsymbol{q}^T P \boldsymbol{q}$, with the following constraints:

$$\big(\boldsymbol{x}^{\mathcal{S}}_{\mathcal{C}_i}, \boldsymbol{\theta}^{\mathcal{S}}_{\mathcal{C}_i}\big) = \big(\check{\boldsymbol{x}}^{\mathcal{S}}_{\mathcal{C}_i}, \check{\boldsymbol{\theta}}^{\mathcal{S}}_{\mathcal{C}_i}\big),$$
$$\big(\boldsymbol{x}^{\mathcal{S}}_{\mathcal{C}_n}, \boldsymbol{\theta}^{\mathcal{S}}_{\mathcal{C}_n}\big) = \Big(\hat{\boldsymbol{x}}^{\mathcal{S}}_{\mathcal{C}_n}, \hat{\boldsymbol{\theta}}^{\mathcal{S}}_{\mathcal{C}_n}\Big) \forall n \in \mathcal{K}_0.$$

As a result of this pose-graph optimization, any discontinuity at the merge location between the prior and current keyframes is smoothed out, as shown in Fig. 6.3. Subsequently, the map points as seen in the current keyframes are also adjusted in accordance to the pose-graph optimization. Finally, the duplicated map points near the merge location are fused together. As a result of shared feature matches between the current and prior keyframes, the updated covisibility window for K_i includes keyframes from prior session(s).

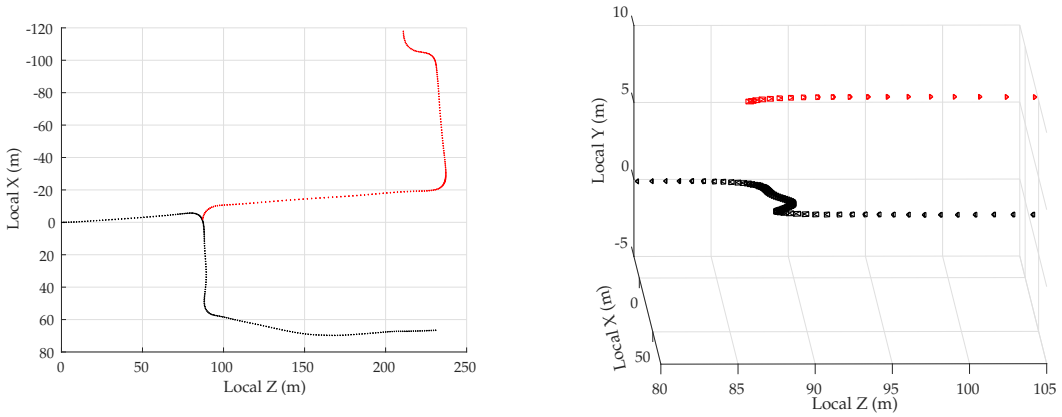


Figure 6.2: GEOSLAM trajectories at the instant when a merge has been detected and verified. The cameras colored black are keyframes from a prior map, and those colored red are from the current session. (**a**) Top view of the trajectories. (**b**) View from 5° elevation showing a discontinuity in the vertical component.

The merged map is then optimized in a windowed BA, this time with feature point coordinates and GNSS positions as measurements. Note that this is a joint windowed BA with both current and prior keyframes and map points. As a result, both the current and prior states are appropriately adjusted based on the number and covariance of the feature point and GNSS measurements. The result of the map merging routine is shown in Fig. 6.4.

6.7 Empirical Results

To validate the results obtained in the above analyses, GNSS and visual data were collected in a moderate urban area north of the University of Texas at Austin campus in Austin, TX. This section presents the results from

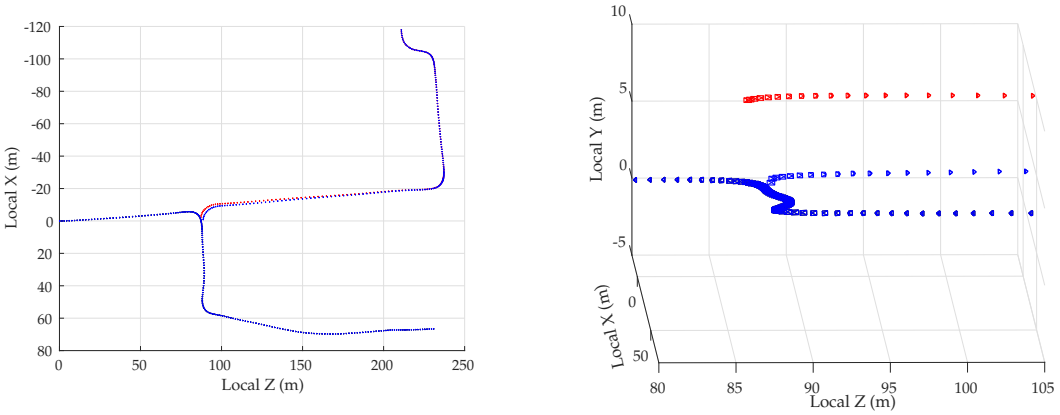


Figure 6.3: GEOSLAM trajectories post-pose-graph optimization (in blue), overlaid on the corresponding (black and red) trajectories from Fig. 6.2. All keyframes are colored blue at this stage since prior and current keyframes are now connected. (**a**) Top view of the trajectories. Note that the discontinuity at the merge location is smoothly distributed across N_m levels of covisibility in the current session, and that the keyframe poses from the prior map are unchanged at this stage. (**b**) View from 5° elevation. Keyframes from the current trajectory have been adjusted to remove the discontinuity, blue and black keyframes exactly overlap. Not shown: the corresponding map points in the current session are also adjusted to match the updated keyframe poses.

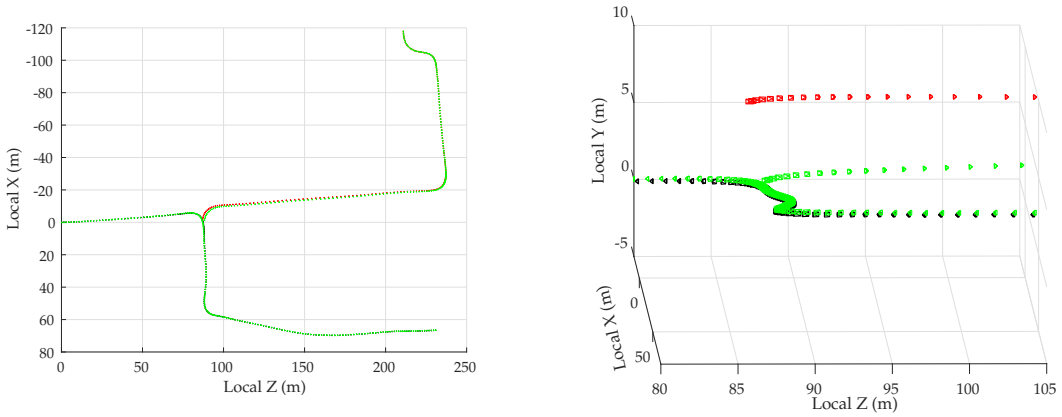


Figure 6.4: GEOSLAM trajectories after joint BA of current and prior keyframes (in green), overlaid on the corresponding (black and red) trajectories from Fig. 6.2. (**a**) Top view of the trajectories. Note that both the current and prior keyframes (and map points, not shown) have been adjusted to optimally minimize the BA cost function over N_m levels of covisibility. (**b**) View from 5° elevation.

GEOSLAM's multi-session GNSS-aided-visual mapping.

6.7.1 Sensor Platform & Test Route

The sensor platform and test route considered in this chapter are identical to ones described in Sec. 5.5. In fact, the data used here are a subset of the data analyzed in Sec. 5.5. Multi-session mapping with GEOSLAM is performed over eight laps/sessions of data from the four data collection campaigns conducted on June 3, 2018 and June 4, 2018.

Imagery collected over the four June 2018 campaigns exhibits appreciable visual diversity, offering a real-world challenge to multi-session GEOSLAM operation. Figs. 6.5a,b show the variation in lighting and visual features be-

Figure 6.5: Different visual conditions on two days of data collection. (**a**) An image captured on the first day of data collection. Note the sharp shadows and absence of parked cars. (**b**) An image captured on the second day of data collection. Note the absence of sharp shadows and complete blockage of curb due to parked cars.

tween the data collected on June 3, 2018 and June 4, 2018.

6.7.2 Multi-Session Mapping Results

GEOSLAM processed two laps/sessions of data from each of the four campaigns conducted on June 4 and 5, fusing the visual data from the captured images with the double-differenced pseudorange-based position estimates of the primary antenna. Fig. 6.6 summarizes the result from GEOSLAM's multi-session GNSS-aided-visual SLAM. The black data points denote the difference between the ground truth trajectory of the primary GNSS antenna and GEOSLAM's estimate of the same in local east, north, and up directions for all eight sessions. The gray data points denote the difference between the ground truth trajectory of the primary GNSS antenna and the coincident double-differenced pseudorange-based estimate of the same for all eight sessions.

As one might expect, the error in GEOSLAM's estimate of the antenna position is approximately the same as the average double-differenced pseudorange-based error over eight sessions. Furthermore, due to the approximately zero-mean nature of the double-differenced pseudorange-based estimates, the GEOSLAM estimate of the trajectory is within 50 cm of the truth trajectory in the horizontal plane. Note that the error in GEOSLAM's position estimate is highly repeatable over eight different sessions, so much so that it appears there is a single black trace in Fig. 6.6, while in truth eight independent traces were plotted. This indicates that (i) the localization of the

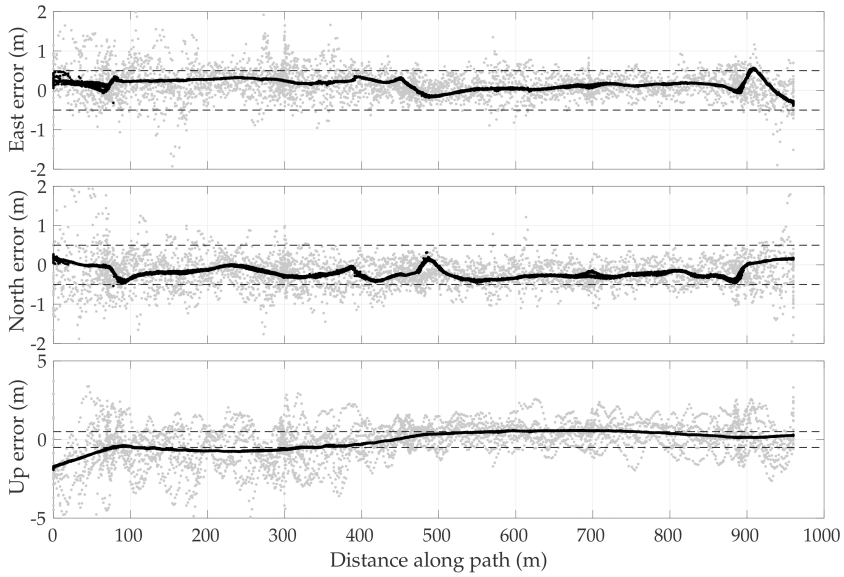


Figure 6.6: Errors in GEOSLAM's estimate of the primary antenna position (in black) with respect to ground truth in the east, north, and up directions for eight mapping sessions from four different data collection campaigns. The errors in double-differenced pseudorange-based primary antenna position estimates for each of the eight sessions, fed as measurements to GEOSLAM, are plotted in gray for reference. Dashed reference lines are drawn at ±50 cm.

vehicle within the visual map was highly precise: GEOSLAM made the same errors with respect to ground truth over eight different sessions; and (ii) the visual map was merged across eight sessions from four different campaigns: if the maps from any two campaigns were not merged through visual matching of features, then the GNSS position estimate for a keyframe from one campaign would not affect another keyframe from a different campaign since they would not be covisible, and thus the eight black traces in Fig. 6.6 would not overlap.

6.8 Conclusion

A globally-referenced electro-optical SLAM pipeline, termed GEOSLAM, has been presented. Notably, this chapter details the techniques that GEOSLAM invokes to smoothly transition between unmapped and previously-mapped regions, including initialization for GNSS-aided SLAM and crowd-sourced map merging. GEOSLAM enables multi-agent collaborative mapping by storing and rendering its map in a global frame of reference. GEOSLAM is demonstrated to achieve sub-50-cm horizontal localization accuracy in a moderate urban environment by incorporating code-phase-based GNSS position estimates in the visual SLAM framework and jointly optimizing maps merged across time-separated sessions.

Chapter 7

Conclusions & Future Work

Four contributions have been presented in support of robust and secure PNT for automated systems. First, a fundamental theory for provably-secure clock synchronization was established. In contrast to prior work in this field, the security conditions identified were shown to be both necessary and sufficient for provably-secure clock synchronization. Second, a three-year study of world-wide GPS interference was presented. Using data from a GNSS receiver on the International Space Station, three major hotspots of persistent and ongoing GNSS interference were detected. This work was a part of the first-ever space-based survey of GNSS interference. Third, a robust, all-weather radar-based ground vehicle positioning system was developed. The proposed system relies on sensors that are available on automated vehicles and are insensitive to lighting and inclement weather: automotive radars, a low-cost IMU, and GNSS. Remarkably, it was shown that, given a prior radar map, these off-the-shelf all-weather automotive sensors maintained sub-50-cm horizontal position accuracy during 60 min of GNSS-denied driving in downtown Austin, TX. Fourth, an analysis and demonstration of the feasibility of crowd-sourced digital mapping for automated vehicles was presented. In an experiment involving multiple laps of a 1 km semi-urban route, it was shown that low-cost consumer

vehicles can incrementally improve the accuracy of a jointly-optimized digital map over time enough to support sub-lane-level positioning in a global frame of reference.

7.1 Future Work

The conditions for secure clock synchronization proposed in Chap. 2, while shown to be necessary and sufficient, are difficult to satisfy under certain scenarios. For example, an alert limit of 1 µs demands that the RTT between `A` and `B` be known to well within 1 µs *a priori*, and that the path between `A` and `B` be irreducible to well within 1 µs. Such conditions are nearly impossible to meet over the Internet, but may be met in a wireless sensor network or a dedicated local area network with reasonable care. Accordingly, it would be instructive to conduct a survey of the achievable alert limit under a variety of synchronization network scenarios.

The global GPS interference survey of Chap. 3 was conducted on the ISS with an aft-pointing antenna flanked by moving solar panels. As such, the setup was only sufficient for a proof-of-concept. An ideal LEO interference probe would be equipped with two GNSS antennas: one pointed towards zenith to track the authentic signals (while attenuating the interfering signals), and the other pointed towards nadir to receive interfering signals from the ground. Data from such probe(s) should dramatically improve the sensitivity and resolution with which global GNSS interference sources can be characterized.

The CINR-based detection technique described in this dissertation only provides a coarse strength- and location-characterization of the interference source with a single probe in LEO. Similarly, the Doppler-based source localization technique of [106] works with a single LEO probe so long as a Doppler time-history can be obtained from the interfering signal. In general, for signals from which no carrier can be isolated, multiple synchronized LEO-based sensors must be deployed with TDOA and FDOA techniques to infer the source's location [20, 21]. A constellation of synchronized dual-antenna probes would be the ideal next step for this dissertation's second contribution.

Space-based interference monitoring and countermeasures shall inevitably be met with counter-countermeasures by interference sources that pretend to be situated at a location other than their actual coordinates. For example, with a sufficiently-accurate prediction of satellite's position and velocity, an interference source may generate a Doppler signature that leads to incorrect geolocation by the probing satellite. Future work must address such adversaries, e.g., with use of directional antennas and TDOA techniques.

This dissertation has presented a remarkable proof-of-concept for all-weather sub-50-cm radar-inertial positioning. The field experiments presented in Chap. 4 evaluate the robustness of the proposed technique by introducing deliberate variation in traffic and parking patterns between the mapping and localization sessions. Nevertheless, further experimentation is required to assess the limits of structual variation in the radar environment that can be handled by the current method. Additionally, while all sensors involved

in the proposed pipeline are weather-resistant in theory, it may be instructive to evaluate the system's performance during and/or after heavy rain and snowfall.

The inverse sensor model applied in Chap. 4 is pessimistic and avoids explicit characterization of the detection and clutter distributions for low-cost automotive radars. While convenient to avoid, such characterization may further improve localization performance, especially if combined with estimation frameworks that directly incorporate these parameters, such as PHD-SLAM [101] with extended target tracking, or its discrete approximation based on the bin occupancy filter [51].

While the localization pipeline in Chap. 4 already estimates the intrinsic and extrinsic calibration parameters for many sensors during periods of CDGNSS availability, in some cases it may be necessary to estimate even more calibration parameters to achieve maximum accuracy. In particular, intrinsic calibration of automotive radars, i.e., estimation of biases in the reported range, range rate, and bearing, may lead to improved mapping and localization accuracy.

Chaps. 5 and 6 explore the limit of mapping and localization accuracy with standard low-cost code-phase GNSS. Simulation modeling and field data collected in a moderately-urban environment indicate that sub-50-cm accuracy is achievable with vision-based GNSS-aided SLAM. While the field data analyzed in this dissertation support the simulation results, a much larger and diverse dataset (with cm-accurate ground truth reference) would be necessary

to claim that the simulation results are reprsentative of the in-field performance of standard GNSS receivers.

GEOSLAM, the GNSS-aided visual SLAM pipeline developed in Chap. 6, jointly optimizes camera and GNSS data from multiple sessions to incrementally improve the mapping and localization accuracy. An obvious drawback of such a design is that the required computational resources grow unbounded as the number of optimization sessions increase over time. While the batch approach of GEOSLAM is important to represent the time-correlated errors between multiple sessions, in practice it would be necessary to develop a windowed or recursive approach that bounds the required computation.

Appendices

Appendix A

Appendix to Chapter 4

A.1 Partial Derivatives

A.1.1 Linearized Forward Dynamics

A few block components of F_k and G_k from (4.7) and (4.8) are listed below.

$$\frac{\partial \delta \boldsymbol{p}_{k+1}^{\mathbf{n}}}{\partial \boldsymbol{\eta}_k^{\mathbf{n}}}\Big|_{\substack{\delta \boldsymbol{x}_k = 0 \\ \boldsymbol{w}_k = 0}} = -\frac{T^2}{2}\left(\tilde{R}_k^{\mathbf{nb}}\left[\boldsymbol{z}_{a,k}^{\mathbf{b}} - \tilde{\boldsymbol{b}}_{a,k}^{\mathbf{b}}\right]_\times\right)$$
$$\frac{\partial \delta \boldsymbol{p}_{k+1}^{\mathbf{n}}}{\partial \delta \boldsymbol{b}_{a,k}^{\mathbf{b}}}\Big|_{\substack{\delta \boldsymbol{x}_k = 0 \\ \boldsymbol{w}_k = 0}} = -\frac{T^2}{2}\tilde{R}_k^{\mathbf{nb}}$$

The partial derivates of $\delta \boldsymbol{v}_{k+1}^{\mathbf{n}}$ with respect to $\delta \boldsymbol{x}_k$ follow similarly.

$$\begin{aligned}
\frac{\partial \boldsymbol{\eta}_{k+1}^{\mathbf{n}}}{\partial \boldsymbol{\eta}_k^{\mathbf{n}}}\Big|_{\substack{\delta \boldsymbol{x}_k = 0 \\ \boldsymbol{w}_k = 0}} &\approx I_{3\times 3} \\
\frac{\partial \boldsymbol{\eta}_{k+1}^{\mathbf{n}}}{\partial \delta \boldsymbol{b}_{\omega,k}^{\mathbf{b}}}\Big|_{\substack{\delta \boldsymbol{x}_k = 0 \\ \boldsymbol{w}_k = 0}} &\approx -T\tilde{R}_{k+1}^{\mathbf{nb}} \mathrm{J}_r\left(\frac{T}{2}\left(\boldsymbol{z}_{\omega,k}^{\mathbf{b}} - \tilde{\boldsymbol{b}}_{\omega,k}^{\mathbf{b}} - \tilde{R}_k^{\mathbf{bn}}\boldsymbol{\omega}_{\mathbf{e}}^{\mathbf{n}}\right)\right) \\
&\approx -T\tilde{R}_{k+1}^{\mathbf{nb}}
\end{aligned}$$

where

$$\mathrm{J}_r(\boldsymbol{\theta}) = I_{3\times 3} - \frac{1-\cos\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^2}[\boldsymbol{\theta}]_\times + \frac{\|\boldsymbol{\theta}\| - \sin\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^3}[\boldsymbol{\theta}]_\times^2$$

is the right Jacobian of $SO(3)$ [145].

$$\begin{aligned}
\frac{\partial \boldsymbol{\eta}_{k+1}^{\mathbf{n}}}{\partial \boldsymbol{w}_{\omega,k}}\Big|_{\substack{\delta \boldsymbol{x}_k = 0 \\ \boldsymbol{w}_k = 0}} &\approx T\tilde{R}_{k+1}^{\mathbf{nb}} \mathrm{J}_r\left(\frac{T}{2}\left(\boldsymbol{z}_{\omega,k}^{\mathbf{b}} - \tilde{\boldsymbol{b}}_{\omega,k}^{\mathbf{b}} - \tilde{R}_k^{\mathbf{bn}}\boldsymbol{\omega}_{\mathbf{e}}^{\mathbf{n}}\right)\right) \\
&\approx T\tilde{R}_{k+1}^{\mathbf{nb}}
\end{aligned}$$

A.1.2 Linearized Measurement Models

The partial derivative of the measurement $\boldsymbol{z}^{\mathbf{n}}_{\mathbf{a}_i,k}$ from (4.9) can be expressed as

$$\frac{\partial \boldsymbol{z}^{\mathbf{n}}_{\mathbf{a}_i,k}}{\partial \delta \boldsymbol{x}_k}\Big|_{\substack{\delta \boldsymbol{x}_k = 0 \\ \boldsymbol{e}_{\mathbf{a}_i,k}=0}} = \frac{\partial \boldsymbol{z}^{\mathbf{n}}_{\mathbf{a}_i,k}}{\partial \boldsymbol{x}_k}\Big|_{\substack{\delta \boldsymbol{x}_k = 0 \\ \boldsymbol{e}_{\mathbf{a}_i,k}=0}} \cdot \frac{\partial \boldsymbol{x}_k}{\partial \delta \boldsymbol{x}_k}\Big|_{\substack{\delta \boldsymbol{x}_k = 0 \\ \boldsymbol{e}_{\mathbf{a}_i,k}=0}}$$

where the non-trivial block matrices are as follows:

$$\frac{\partial \boldsymbol{z}^{\mathbf{n}}_{\mathbf{a}_i,k}}{\partial \boldsymbol{q}^{\mathbf{nb}}_k}\Big|_{\substack{\delta \boldsymbol{x}_k = 0 \\ \boldsymbol{e}_{\mathbf{a}_i,k}=0}} = \frac{\partial \big(\boldsymbol{q}^{\mathbf{nb}}_k \odot R^{\mathbf{bs}} \boldsymbol{p}^{\mathbf{s}}_{\mathbf{ba}_i} \odot \boldsymbol{q}^{\mathbf{bn}}_k\big)}{\partial \boldsymbol{q}^{\mathbf{nb}}_k}$$

$$\frac{\partial \boldsymbol{q}^{\mathbf{nb}}_k}{\partial \boldsymbol{\eta}^{\mathbf{n}}_k}\Big|_{\substack{\delta \boldsymbol{x}_k = 0 \\ \boldsymbol{e}_{\mathbf{a}_i,k}=0}} = \frac{1}{2}\begin{bmatrix} -q_x & -q_y & -q_z \\ q_w & q_z & -q_y \\ -q_z & q_w & q_x \\ q_y & -q_x & q_w \end{bmatrix}$$

with $\tilde{\boldsymbol{q}}^{\mathbf{nb}}_k = [q_w, q_x, q_y, q_z]$. The expression for derivative of the rotation with respect to the quaternion can be found in [145, Sec. 4.3.2].

For the radar range rate measurement $\boldsymbol{z}^{\mathbf{r}_i}_{\mathbf{r}_i,k}$

$$\frac{\partial \boldsymbol{z}^{\mathbf{r}_i}_{\mathbf{r}_i,k}}{\partial \boldsymbol{v}^{\mathbf{n}}_k} = R^{\mathbf{r}_i\mathbf{s}} R^{\mathbf{sb}} \tilde{R}^{\mathbf{bn}}_k$$

$$\frac{\partial \boldsymbol{z}^{\mathbf{r}_i}_{\mathbf{r}_i,k}}{\partial \boldsymbol{q}^{\mathbf{nb}}_k} = R^{\mathbf{r}_i\mathbf{s}} R^{\mathbf{sb}} \frac{\partial \big(\boldsymbol{q}^{\mathbf{bn}}_k \odot \tilde{\boldsymbol{v}}^{\mathbf{n}}_k \odot \boldsymbol{q}^{\mathbf{nb}}_k\big)}{\partial \boldsymbol{q}^{\mathbf{nb}}_k}$$

$$\frac{\partial \boldsymbol{z}^{\mathbf{r}_i}_{\mathbf{r}_i,k}}{\partial \boldsymbol{b}^{\mathbf{b}}_{\omega,k}} = -R^{\mathbf{r}_i\mathbf{s}} R^{\mathbf{sb}} \big[R^{\mathbf{bs}} \boldsymbol{p}^{\mathbf{s}}_{\mathbf{br}_i}\big]_\times$$

The partial derivatives of $\boldsymbol{z}^{\mathbf{v}}_{\text{nhc},k}$ and $\boldsymbol{z}^{\mathbf{v}}_{\text{zupt},k}$ follow similarly.

A.2 Nonlinear Error-State Rauch-Tung-Striebel Smoother

The conventional expression for the extended Rauch-Tung-Striebel (RTS) smoother is given as [132, Chap. 9]

$$\begin{aligned}\boldsymbol{x}_k^\star &= \hat{\boldsymbol{x}}_k + C_k\big(\boldsymbol{x}_{k+1}^\star - f_k(\hat{\boldsymbol{x}}_k)\big)\\ P_k^\star &= P_k + C_k\big(P_{k+1}^\star - F_k P_k F_k^\top - G_k Q_k G_k^\top\big)C_k^\top\end{aligned}$$

with

$$C_k = P_k F_k^\top \big(F_k P_k F_k^\top + G_k Q_k G_k^\top\big)^{-1}$$

where * indicates the smoothed estimate and $\hat{}$ indicates the filtered estimate. This expression is derived by linearizing the dynamics at the *filtered* state estimate during the *backward smoothing* pass.

In contrast, this paper prefers to linearize the dynamics at the predicted smoothed estimate $\bar{\boldsymbol{x}}_k^\star$ instead

$$\bar{\boldsymbol{x}}_k^\star \triangleq f_k^{-1}\big(\boldsymbol{x}_{k+1}^\star, \boldsymbol{u}_k, \boldsymbol{0}\big)$$

This formulation results in a similar but slightly modified expression for the extended RTS smoother

$$\begin{aligned}\boldsymbol{x}_k^\star &= \hat{\boldsymbol{x}}_k + C_k^\star F_k^\star(\bar{\boldsymbol{x}}_k^\star - \hat{\boldsymbol{x}}_k)\\ P_k^\star &= P_k + C_k^\star\big(P_{k+1}^\star - F_k^\star P_k F_k^{\star\top} - G_k^\star Q_k G_k^{\star\top}\big)C_k^{\star\top}\end{aligned}$$

with

$$C_k^\star = P_k F_k^{\star\top}\big(F_k^\star P_k F_k^{\star\top} + G_k^\star Q_k G_k^{\star\top}\big)^{-1}$$

where $F_k^\star$ and $G_k^\star$ denote linearized forward dynamics around $\bar{\boldsymbol{x}}_k^\star$.

Bibliography

[1] GPS satellite ephemerides / satellite & station clocks. http://www.igs.org/products. Accessed: 2018-03-30.

[2] Dennis M Akos. Who's afraid of the spoofer? GPS/GNSS spoofing detection via automatic gain control (AGC). *Navigation, Journal of the Institute of Navigation*, 59(4):281–290, 2012.

[3] Evangelos Bitsikas AlaDarabseh and Brice Tedongmo. Detecting GPS jamming incidents in OpenSky data. In *Proceedings of the 7th OpenSky Workshop*, volume 67, pages 97–108, 2019.

[4] David W Allan and Marc Abbott Weiss. *Accurate time and frequency transfer during common-view of a GPS satellite*. Electronic Industries Association, 1980.

[5] Yasin Almalioglu, Mehmet Turan, Chris Xiaoxuan Lu, Niki Trigoni, and Andrew Markham. Milli-RIO: Ego-motion estimation with low-cost millimetre-wave radar. *arXiv preprint arXiv:1909.05774*, 2019.

[6] L. A. A. Andersson and J. Nygards. C-SAM: Multi-robot slam using square root information smoothing. In *2008 IEEE International Conference on Robotics and Automation*, pages 2798–2805, May 2008.

[7] James J. Angel. Impact of special relativity on securities regulation.

The Future of Computer Trading in Financial Markets. Foresight Driver Review–DR 15, April 2011.

[8] James J Angel. When finance meets physics: The impact of the speed of light on financial markets and their regulation. *Financial Review*, 2(49):271–281, 2014.

[9] Robert Annessi, Joachim Fabini, and Tanja Zseby. SecureTime: Secure multicast time synchronization. *arXiv preprint arXiv:1705.10669*, 2017.

[10] Bernhard Martin Aumayer. *Ultra-tightly Coupled Vision/GNSS for Automotive Applications*. PhD thesis, University of Calgary, 2016.

[11] Yaakov Bar-Shalom, X. Rong Li, and Thiagalingam Kirubarajan. *Estimation with Applications to Tracking and Navigation*. John Wiley and Sons, New York, 2001.

[12] Dan Barnes and Ingmar Posner. Under the radar: Learning to predict robust keypoints for odometry estimation and metric localisation in radar. *arXiv preprint arXiv:2001.10789*, 2020.

[13] Dan Barnes, Rob Weston, and Ingmar Posner. Masking by moving: Learning distraction-free radar odometry from pose information. *arXiv preprint arXiv:1909.03752*, 2019.

[14] Jeremy Barra, Suzanne Lesecq, Mykhailo Zarudniev, Olivier Debicki, Nicolas Mareau, and Laurent Ouvry. Localization system in GPS-denied environments using radar and IMU measurements: Application

to a smart white cane. In *2019 18th European Control Conference (ECC)*, pages 1201–1206. IEEE, 2019.

[15] A Bauch, P Hetzel, and D Piester. Time and frequency dissemination with DCF77: From 1959 to 2009 and beyond. *PTB-Mitteilungen*, 119(3):3–26, 2009.

[16] BBC. What would the world do without GPS? https://bbc.in/3jedl4q. Accessed 2020-10-19.

[17] Mihir Bellare and Chanathip Namprempre. Authenticated encryption: Relations among notions and analysis of the generic composition paradigm. *Advances in Cryptology–ASIACRYPT 2000*, pages 531–545, 2000.

[18] Charles H Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. 1984.

[19] Sriramya Bhamidipati, Yuting Ng, and Grace Xingxin Gao. Multi-receiver GPS-based direct time estimation for PMUs. In *Proceedings of the 29th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2016), Portland, OR*, 2016.

[20] Jahshan Bhatti. *Sensor Deception Detection and Radio-Frequency Emitter Localization*. PhD thesis, The University of Texas at Austin, Aug. 2015.

[21] Jahshan Bhatti and Todd E Humphreys. Hostile control of ships via

false GPS signals: Demonstration and detection. *Navigation*, 64(1):51–66, 2017.

[22] Johannes Boehm, Birgit Werl, and Harald Schuh. Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for Medium-Range Weather Forecasts operational analysis data. *Journal of Geophysical Research: Solid Earth*, 111(B2), 2006.

[23] Johannes Böhm, Gregor Möller, Michael Schindelegger, Gregory Pain, and Robert Weber. Development of an improved empirical model for slant delays in the troposphere (GPT2w). *GPS Solutions*, 19(3):433–441, 2015.

[24] Johannes Böhm, Arthur Niell, Paul Tregoning, and Harald Schuh. Global mapping function (GMF): A new empirical mapping function based on numerical weather model data. *Geophysical Research Letters*, 33(7), 2006.

[25] Daniele Borio. PANOVA tests and their application to GNSS spoofing detection. *IEEE Transactions on Aerospace and Electronic Systems*, 49(1):381–394, Jan. 2013.

[26] Michael S. Braasch. *Springer Handbook of Global Navigation Satellite Systems*, chapter Multipath, pages 443–468. Springer, 2017.

[27] Ben Brimelow. General reveals that US aircraft are being 'disabled' in Syria the 'most aggressive' electronic warfare environment on Earth, April 2018. https://goo.gl/9B2Nf4.

[28] Ali Broumandan, Sandy Kennedy, and John Schleppe. Demonstration of a multi-layer spoofing detection implemented in a high precision gnss receiver. In *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pages 538–547. IEEE, 2020.

[29] Mitch Bryson and Salah Sukkarieh. A comparison of feature and pose-based mapping using vision, inertial and GPS on a UAV. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, San Francisco, CA, Sept. 2011. IEEE.

[30] C4ADS. Above us only stars: Exposing GPS spoofing in Russia and Syria, April 2019. https://c4ads.org/reports.

[31] C4ISR. Army to award contract for GPS alternative by end of September. https://bit.ly/3dOx0a8. Accessed 2020-10-19.

[32] Jonas Callmer, David Törnqvist, Fredrik Gustafsson, Henrik Svensson, and Pelle Carlbom. Radar SLAM using visual features. *EURASIP Journal on Advances in Signal Processing*, 2011(1):71, 2011.

[33] Sarah H Cen and Paul Newman. Precise ego-motion estimation with millimeter-wave radar under diverse and challenging conditions. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1–8. IEEE, 2018.

[34] Sarah H Cen and Paul Newman. Radar-only ego-motion estimation in difficult settings via graph matching. *arXiv preprint arXiv:1904.11476*, 2019.

[35] Dmitry Chetverikov, Dmitry Svirko, Dmitry Stepanov, and Pavel Krsek. The trimmed iterative closest point algorithm. In *Object recognition supported by user interaction for service robots*, volume 3, pages 545–548. IEEE, 2002.

[36] Kai-Wei Chiang, Guang-Je Tsai, Yu-Hua Li, You Li, and Naser El-Sheimy. Navigation engine design for automated driving using INS/GNSS/3D LiDAR-SLAM and integrity assessment. *Remote Sensing*, 12(10):1564, 2020.

[37] J. Choi, V. Va, N. Gonzalez-Prelcic, R. Daniels, C. R. Bhat, and R. W. Heath. Millimeter-wave vehicular communication to support massive automotive sensing. *IEEE Communications Magazine*, 54(12):160–167, December 2016.

[38] Daniel Chou, Liang Heng, and GX Gao. Robust GPS-based timing for phasor measurement units: A position-information-aided approach. In *Proceedings of the ION GNSS+ Meeting*, 2014.

[39] Tianxing Chu, Ningyan Guo, Staffan Backén, and Dennis Akos. Monocular camera/IMU/GNSS integration for ground vehicle navigation in challenging GNSS environments. *Sensors*, 12(3):3162–3185, 2012.

[40] Mihaela-Simona Circiu, Michael Meurer, Michael Felux, Daniel Gerbeth, Steffen Thölert, Mariano Vergara, Christoph Enneking, Mateo Sgammini, Samuel Pullen, and Felix Antreich. Evaluation of GPS L5 and Galileo E1 and E5a performance for future multifrequency and multi-

constellation GBAS. *Navigation*, 64(1):149–163, 2017.

[41] Todor Cooklev, John C Eidson, and Afshaneh Pakdaman. An implementation of IEEE 1588 over IEEE 802.11b for synchronization of wireless local area network nodes. *IEEE Transactions on Instrumentation and Measurement*, 56(5):1632–1639, Oct. 2007.

[42] James C Corbett, Jeffrey Dean, Michael Epstein, Andrew Fikes, Christopher Frost, Jeffrey John Furman, Sanjay Ghemawat, Andrey Gubarev, Christopher Heiser, Peter Hochschild, et al. Spanner: Google's globally distributed database. *ACM Transactions on Computer Systems (TOCS)*, 31(3):8, 2013.

[43] John L Crassidis, F Landis Markley, and Yang Cheng. Survey of nonlinear attitude estimation methods. *Journal of guidance control and dynamics*, 30(1):12, 2007.

[44] A. Cunningham, K. M. Wurm, W. Burgard, and F. Dellaert. Fully distributed scalable smoothing and mapping with robust multi-robot data association. In *2012 IEEE International Conference on Robotics and Automation*, pages 1093–1100, May 2012.

[45] O. Dabeer, W. Ding, R. Gowaiker, S. K. Grzechnik, M. J. Lakshman, S. Lee, G. Reitmayr, A. Sharma, K. Somasundaram, R. T. Sukhavasi, and X. Wu. An end-to-end system for crowdsourced 3d maps for autonomous vehicles: The mapping component. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages

634–641, Sept 2017.

[46] Department of Homeland Security. National risk estimate: Risks to U.S. critical infrastructure from Global Positioning System disruptions, November 2012. FOUO: No Public Version Available.

[47] Hendrik Deusch, Stephan Reuter, and Klaus Dietmayer. The labeled multi-Bernoulli SLAM filter. *IEEE Signal Processing Letters*, 22(10):1561–1565, 2015.

[48] Artur K Ekert. Quantum cryptography based on Bell's theorem. *Physical review letters*, 67(6):661, 1991.

[49] Cameron Ellum. Integration of raw GPS measurements into a bundle adjustment. *IAPRS series*, 35(3025), 2006.

[50] Jakob Engel, Thomas Schöps, and Daniel Cremers. LSD-SLAM: Large-scale direct monocular SLAM. In *European Conference on Computer Vision*, pages 834–849. Springer, 2014.

[51] Ozgur Erdinc, Peter Willett, and Yaakov Bar-Shalom. The bin-occupancy filter and its connection to the PHD filters. *IEEE Transactions on Signal Processing*, 57(11):4232–4246, 2009.

[52] Ericsson Blog. 5G is all in the timing. https://bit.ly/2FIdNtW. Accessed 2020-10-19.

[53] David Fajardo, Tsz-Chiu Au, S Waller, Peter Stone, and David Yang. Automated intersection control: Performance of future innovation versus

current traffic signal control. *Transportation Research Record: Journal of the Transportation Research Board*, (2259):223–232, 2011.

[54] Maryam Fatemi, Karl Granström, Lennart Svensson, Francisco JR Ruiz, and Lars Hammarstrand. Poisson multi-bernoulli mapping using Gibbs sampling. *IEEE Transactions on Signal Processing*, 65(11):2814–2827, 2017.

[55] C. Forster, S. Lynen, L. Kneip, and D. Scaramuzza. Collaborative monocular slam with multiple micro aerial vehicles. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3962–3970, Nov 2013.

[56] Wei Gao and Russ Tedrake. FilterReg: Robust and efficient probabilistic point-set registration using gaussian filter and twist parameterization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 11095–11104, 2019.

[57] J. Giri, D. Sun, and R. Avila-Rosales. Wanted: A more intelligent grid. *IEEE Power & Energy*, pages 34–40, April 2009.

[58] Saurabh Godha. Performance evaluation of low cost MEMS-based IMU integrated with GPS for land vehicle navigation application. Master's thesis, The University of Calgary, 2006.

[59] Shafi Goldwasser, Silvio Micali, and Ronald L Rivest. A digital signature scheme secure against adaptive chosen-message attacks. *SIAM Journal on Computing*, 17(2):281–308, 1988.

[60] Joseph C. Grabowski. Personal Privacy Jammers: Locating Jersey PPDs Jamming GBAS Safety-of-Life Signals. *GPS World*, 23(4):28–37, April 2012.

[61] Andy Greenberg. How an entire nation became Russia's test lab for cyberwar, Jun 2017.

[62] Martin Holder, Sven Hellwig, and Hermann Winner. Real-time pose graph SLAM based on radar. In *2019 IEEE Intelligent Vehicles Symposium (IV)*, pages 1145–1151. IEEE, 2019.

[63] Ziyang Hong, Yvan Petillot, and Sen Wang. RadarSLAM: Radar based large-scale SLAM in all weathers. *arXiv preprint arXiv:2005.02198*, 2020.

[64] T. E. Humphreys, M. Murrian, and L. Narula. Low-cost precise vehicular positioning in urban environments. In *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pages 456–471, April 2018.

[65] Todd E. Humphreys. *Springer Handbook of Global Navigation Satellite Systems*, chapter Interference, pages 469–504. Springer, 2017.

[66] Todd E Humphreys, Brent M Ledvina, Mark L Psiaki, Brady W. O'Hanlon, and Paul M Kintner, Jr. Assessing the spoofing threat: Development of a portable GPS civilian spoofer. In *Proceedings of the ION GNSS Meeting*, Savannah, GA, 2008. Institute of Navigation.

[67] Todd E. Humphreys, Matthew J. Murrian, and Lakshay Narula. Deep urban unaided precise Global Navigation Satellite System vehicle positioning. *IEEE Intelligent Transportation Systems Magazine*, 2020.

[68] J. J. Hutton, N. Gopaul, X. Zhang, J. Wang, V. Menon, D. Rieck, A. Kipka, and F. Pastor. Centimeter-level, robust GNSS-aided inertial post-processing for mobile mapping without local reference stations. volume XLI-B3, pages 819–826, Gottingen, 2016. Copernicus GmbH.

[69] V. Indelman, E. Nelson, N. Michael, and F. Dellaert. Multi-robot pose graph localization and data association from unknown initial relative poses via expectation maximization. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 593–600, May 2014.

[70] Bing Jian and Baba C Vemuri. Robust point set registration using Gaussian mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1633–1645, 2010.

[71] John A. Volpe National Transportation Systems Center. Vulnerability assessment of the transportation infrastructure relying on the Global Positioning System, 2001.

[72] Sandy Kennedy, Jason Hamilton, and Hugh Martell. Architecture and system performance of SPAN—NovAtel's GPS/INS solution. In *Position, Location, And Navigation Symposium, 2006 IEEE/ION*, page 266. IEEE, 2006.

[73] Victor L Knoop, Peter F de Bakker, Christian CJM Tiberius, and Bart

van Arem. Lane determination with GPS precise point positioning. *IEEE Transactions on Intelligent Transportation Systems*, 18(9):2503–2513, 2017.

[74] Manon Kok, Jeroen D Hol, and Thomas B Schön. Using inertial sensors for position and orientation estimation. *arXiv preprint arXiv:1704.06053*, 2017.

[75] Andrew Kramer, Carl Stahoviak, Angel Santamaria-Navarro, Ali-Akbar Agha-Mohammadi, and Christoffer Heckman. Radar-inertial ego-velocity estimation for visually degraded environments. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020.

[76] H. Kume, T. Taketomi, T. Sato, and N. Yokoya. Extrinsic camera parameter estimation using video images and GPS considering GPS positioning accuracy. In *2010 20th International Conference on Pattern Recognition*, pages 3923–3926, Aug 2010.

[77] Daniel LaChapelle, Todd E. Humphreys, Lakshay Narula, Peter A. Iannucci, and Ehsan Moradi-Pari. Automotive collision risk estimation under cooperative sensing. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Barcelona, Spain, 2020.

[78] Anh Quan Le and Christian Tiberius. Single-frequency precise point positioning with optimal filtering. *GPS Solutions*, 11(1):61–69, Jan. 2007.

[79] Andreas Lehner and Alexander Steingass. A novel channel model for land mobile satellite navigation. In *Proceedings of the ION GNSS Meeting*, pages 13–16, 2005.

[80] Andreas Lehner and Alexander Steingass. Technical note on the Land Mobile Satellite Channel Model - Interface Control Document, May 2008.

[81] Stefan Leutenegger, Simon Lynen, Michael Bosse, Roland Siegwart, and Paul Furgale. Keyframe-based visual–inertial odometry using nonlinear optimization. *The International Journal of Robotics Research*, 34(3):314–334, 2015.

[82] Jesse Levinson, Michael Montemerlo, and Sebastian Thrun. Map-based precision vehicle localization in urban environments. In *Robotics: Science and Systems*, volume 4, page 1. Citeseer, 2007.

[83] Jesse Levinson and Sebastian Thrun. Robust vehicle localization in urban environments using probabilistic maps. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pages 4372–4378. IEEE, 2010.

[84] M. Lhuillier. Fusion of GPS and structure-from-motion using constrained bundle adjustments. In *CVPR 2011*, pages 3025–3032, June 2011.

[85] M. Lhuillier. Incremental fusion of structure-from-motion and GPS using constrained bundle adjustments. *IEEE Transactions on Pattern*

Analysis and Machine Intelligence, 34(12):2489–2495, Dec 2012.

[86] Rongbing Li, Jianye Liu, Ling Zhang, and Yijun Hang. Lidar/mems imu integrated navigation (slam) method for a small uav in indoor environments. In *2014 DGON Inertial Sensors and Systems (ISS)*, pages 1–15. IEEE, 2014.

[87] E. Glenn Lightsey, Todd E. Humphreys, Jahshan A. Bhatti, Andrew J. Joplin, Brady W. O'Hanlon, and Steven P. Powell. Demonstration of a space capable miniature dual frequency GNSS receiver. *Navigation, Journal of the Institute of Navigation*, 61(1):53–64, 2014.

[88] LORD Sensing MicroStrain. 3DM-GX5-25 Attitude and Heading Reference System. https://bit.ly/32CKIaO. Accessed 2020-08-31.

[89] David G Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[90] Ronald PS Mahler. Multitarget Bayes filtering via first-order multitarget moments. *IEEE Transactions on Aerospace and Electronic systems*, 39(4):1152–1178, 2003.

[91] Aneeq Mahmood, Georg Gaderer, Henning Trsek, Stefan Schwalowsky, and Nikolaus Kerö. Towards high accuracy in IEEE 802.11 based clock synchronization using PTP. In *Precision Clock Synchronization for Measurement Control and Communication (ISPCS), 2011 International IEEE Symposium on*, pages 13–18. IEEE, 2011.

[92] K. E. Martin and et al. Exploring the IEEE standard C37.118–2005 synchrophasors for power systems. *IEEE Transactions on Power Delivery*, 23(4):1805–1811, Oct. 2008.

[93] Jules G McNeff. The global positioning system. *IEEE Transactions on Microwave Theory and Techniques*, 50(3):645–652, 2002.

[94] Ralph C Merkle. Secure communications over insecure channels. *Communications of the ACM*, 21(4):294–299, 1978.

[95] Michael Meurer, Andriy Konovaltsev, Manuel Cuntz, and Christian Hättich. Robust joint multi-antenna spoofing detection and attitude estimation using direction assisted multiple hypotheses RAIM. In *Proceedings of the 25th Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2012)*. ION, 2012.

[96] Pratap Misra and Per Enge. *Global Positioning System: Signals, Measurements, and Performance*. Ganga-Jumana Press, Lincoln, Massachusetts, revised second edition, 2012.

[97] Tal Mizrahi. A game theoretic analysis of delay attacks against time synchronization protocols. In *Precision Clock Synchronization for Measurement Control and Communication (ISPCS), 2012 International IEEE Symposium on*, pages 1–6. IEEE, 2012.

[98] Bassam Moussa, Mourad Debbabi, and Chadi Assi. A detection and mitigation model for PTP delay attack in an IEC 61850 substation. *IEEE Transactions on Smart Grid*, 2016.

[99] Peter Mühlfellner, Mathias Bürki, Michael Bosse, Wojciech Derendarz, Roland Philippsen, and Paul Furgale. Summary maps for lifelong visual localization. *Journal of Field Robotics*, 33(5):561–590, 2016.

[100] Marius Muja and David G Lowe. Fast approximate nearest neighbors with automatic algorithm configuration. *VISAPP (1)*, 2(331-340):2, 2009.

[101] John Mullane, Ba-Ngu Vo, Martin D Adams, and Ba-Tuong Vo. A random-finite-set approach to Bayesian SLAM. *IEEE Transactions on Robotics*, 27(2):268–282, 2011.

[102] Raul Mur-Artal, Jose Maria Martinez Montiel, and Juan D Tardos. Orb-slam: a versatile and accurate monocular slam system. *IEEE Transactions on Robotics*, 31(5):1147–1163, 2015.

[103] Raul Mur-Artal and Juan D Tardos. ORB-SLAM2: an open-source SLAM system for monocular, stereo and RGB-D cameras. *arXiv preprint arXiv:1610.06475*, 2016.

[104] Raúl Mur-Artal and Juan D Tardós. Visual-inertial monocular SLAM with map reuse. *IEEE Robotics and Automation Letters*, 2(2):796–803, 2017.

[105] Matthew J. Murrian, Collin W. Gonzalez, Todd E. Humphreys, Kenneth M. Pesyna Jr., Daniel P. Shepard, and Andrew J. Kerns. Low-cost precise positioning for automated vehicles. *GPS World*, 27(9):32–39, September 2016.

[106] Matthew J. Murrian, Lakshay Narula, Peter A. Iannucci, Scott Budzien, Brady W. O'Hanlon, Steven P. Powell, and Todd E. Humphreys. GNSS interference monitoring from low Earth orbit. *Navigation, Journal of the Institute of Navigation*, 2020. Submitted for review.

[107] Andriy Myronenko and Xubo Song. Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(12):2262–2275, 2010.

[108] Lakshay Narula and Todd E. Humphreys. Requirements for secure clock synchronization. *IEEE Journal of Selected Topics in Signal Processing*, 12(4):749–762, Aug. 2018.

[109] Lakshay Narula, Peter A Iannucci, and Todd E Humphreys. All-weather sub-50-cm radar-inertial positioning. *Field Robotics*, 2020. Submitted for review.

[110] Lakshay Narula, Peter A Iannucci, and Todd E Humphreys. Automotive-radar-based 50-cm urban positioning. In *Proceedings of the IEEE/ION PLANSx Meeting*, 2020.

[111] Lakshay Narula, Daniel M LaChapelle, Matthew J Murrian, J Michael Wooten, Todd E Humphreys, Jean-Baptiste Lacambre, Elliot de Toldi, and Guirec Morvant. TEX-CUP: The University of Texas Challenge for Urban Positioning. In *Proceedings of the IEEE/ION PLANSx Meeting*, 2020.

[112] Lakshay Narula, Matthew J Murrian, and Todd E Humphreys. Accuracy limits for globally-referenced digital mapping using standard GNSS. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 3075–3082. IEEE, 2018.

[113] Lakshay Narula, J. Michael Wooten, Matthew J. Murrian, Daniel M. LaChapelle, and Todd E. Humphreys. Accurate collaborative globally-referenced digital mapping with standard GNSS. *Sensors*, 18(8), 2018.

[114] NextGov. White House Council seeks input on plan to invest in alternatives to GPS. https://bit.ly/3o8Vpf1. Accessed 2020-10-19.

[115] Yuting Ng and Grace Xingxin Gao. Robust GPS-based direct time estimation for PMUs. In *Position, Location and Navigation Symposium (PLANS), 2016 IEEE/ION*, pages 472–476. IEEE, 2016.

[116] D. Odijk. *Fast Precise GPS Positioning in the Presence of Ionospheric Delays*. Number no. 52 in Fast precise GPS positioning in the presence of ionospheric delays. NCG, Nederlandse Commissie voor Geodesie, 2002.

[117] A. Perrig, R. Canetti, J.D. Tygar, and D. Song. The TESLA broadcast authentication protocol. *RSA CryptoBytes*, 5(2):2–13, 2002.

[118] Kenneth M. Pesyna, Jr. *Advanced Techniques for Centimeter-Accurate GNSS Positioning on Low-Cost Mobile Platforms*. PhD thesis, The University of Texas at Austin, Dec. 2015.

[119] MG Petovello, ME Cannon, and G Lachapelle. Benefits of using a tactical-grade IMU for high-accuracy positioning. *Navigation, Journal of the Institute of Navigation*, 51(1):1–12, 2004.

[120] AG Phadke, B. Pickett, M. Adamiak, M. Begovic, G. Benmouyal, RO Burnett Jr, TW Cease, J. Goossens, DJ Hansen, M. Kezunovic, et al. Synchronized sampling and phasor measurements for relaying and control. *IEEE Transactions on Power Delivery*, 9(1):442–452, 1994.

[121] N. Piasco, J. Marzat, and M. Sanfourche. Collaborative localization and formation flying using distributed stereo-vision. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1202–1207, May 2016.

[122] M. L. Psiaki, T. E. Humphreys, and B. Stauffer. Attackers can spoof navigation signals without our knowledge. here's how to fight back GPS lies. *IEEE Spectrum*, 53(8):26–53, August 2016.

[123] Mark L Psiaki and Todd E Humphreys. GNSS spoofing and detection. *Proceedings of the IEEE*, 104(6):1258–1270, 2016.

[124] M.L. Psiaki and S. Mohiuddin. Modeling, analysis, and simulation of GPS carrier phase for spacecraft relative navigation. *Journal of Guidance Control and Dynamics*, 30(6):1628, 2007.

[125] Tong Qin, Peiliang Li, and Shaojie Shen. VINS-mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Transactions on Robotics*, 34(4):1004–1020, 2018.

[126] Ramaswamy Ramaswamy, Ning Weng, and Tilman Wolf. Characterizing network processing delay. In *Global Telecommunications Conference, 2004. GLOBECOM'04. IEEE*, volume 3, pages 1629–1634. IEEE, 2004.

[127] B Srinivasa Reddy and Biswanath N Chatterji. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE transactions on image processing*, 5(8):1266–1271, 1996.

[128] Seth Rogers. Creating and evaluating highly accurate maps with probe vehicles. In *Intelligent Transportation Systems, 2000. Proceedings. 2000 IEEE*, pages 125–130. IEEE, 2000.

[129] A. Rovira-Garcia, J. M. Juan, J. Sanz, and G. Gonzlez-Casado. A worldwide ionospheric model for fast precise point positioning. *IEEE Transactions on Geoscience and Remote Sensing*, 53(8):4596–4604, Aug 2015.

[130] Adrià Rovira-Garcia, JM Juan, J Sanz, Guillermo González-Casado, and D Ibáñez. Accuracy of ionospheric models used in GNSS and SBAS: methodology and analysis. *Journal of geodesy*, 90(3):229–240, 2016.

[131] Saeedi Sajad, Trentini Michael, Seto Mae, and Li Howard. Multiple-robot simultaneous localization and mapping: A review. *Journal of Field Robotics*, 33(1):3–46.

[132] Simo Särkkä. *Bayesian filtering and smoothing*, volume 3. Cambridge University Press, 2013.

[133] Bruno M Scherzinger. Precise robust positioning with inertially aided RTK. *Navigation*, 53(2):73–83, 2006.

[134] David Schneider. The microsecond market. In *IEEE Spectrum*, pages 67–71,80–81, June 2012.

[135] Markus Schoen, Markus Horn, Markus Hahn, and Juergen Dickmann. Real-time radar SLAM.

[136] Frank Schuster, Christoph Gustav Keller, Matthias Rapp, Martin Haueis, and Cristóbal Curio. Landmark based radar SLAM using graph optimization. In *Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on*, pages 2559–2564. IEEE, 2016.

[137] Clare Sebastian. Getting lost near the Kremlin? Russia could be GPS spoofing. CNN, December 2 2016.

[138] L Dennis Shapiro. Time synchronization from Loran-C. *IEEE Spectrum*, 8(5):46–55, 1968.

[139] D. P. Shepard, T. E. Humphreys, and A. A. Fansler. Evaluation of the vulnerability of phasor measurement units to GPS spoofing attacks. *International Journal of Critical Infrastructure Protection*, 5(3-4):146–153, 2012.

[140] Daniel P. Shepard, Jahshan A. Bhatti, Todd E. Humphreys, and Aaron A. Fansler. Evaluation of smart grid and civilian UAV vulnerability to GPS spoofing attacks. In *Proceedings of the ION GNSS Meeting*, 2012.

[141] Daniel P. Shepard and Todd E. Humphreys. High-precision globally-referenced position and attitude via a fusion of visual SLAM, carrier-phase-based GPS, and inertial measurements. In *Proceedings of the IEEE/ION PLANS Meeting*, May 2014.

[142] Daniel P. Shepard and Todd E. Humphreys. Scalable sub-decimeter-accurate 3d reconstruction. 2017. In preparation.

[143] Chuang Shi, Shengfeng Gu, Yidong Lou, and Maorong Ge. An improved approach to model ionospheric delays for single-frequency precise point positioning. *Advances in Space Research*, 49(12):1698 – 1708, 2012.

[144] Isaac Skog, Peter Handel, John-Olof Nilsson, and Jouni Rantakokko. Zero-velocity detection–An algorithm evaluation. *IEEE transactions on biomedical engineering*, 57(11):2657–2666, 2010.

[145] Joan Sola. Quaternion kinematics for the error-state Kalman filter. *arXiv preprint arXiv:1711.02508*, 2017.

[146] Olivia Solon. GPS 'spoofers' could be used for high-frequency financial trading fraud. Wired.co.uk, Feb. 2012.

[147] Andrey Soloviev and Donald Venable. Integration of GPS and vision measurements for navigation in GPS challenged environments. In *Proceedings of the IEEE/ION PLANS Meeting*, pages 826–833. IEEE/Institute of Navigation, May 2010.

[148] Olga Sorkine-Hornung and Michael Rabinovich. Least-squares rigid motion using SVD. *Computing*, 1:1, 2017.

[149] Bastian Steder, Giorgio Grisetti, Cyrill Stachniss, and Wolfram Burgard. Visual SLAM for flying vehicles. *IEEE Transactions on Robotics*, 24(5):1088–1093, 2008.

[150] Hauke Strasdat, Andrew J Davison, JM Martìnez Montiel, and Kurt Konolige. Double window optimisation for constant time visual slam. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 2352–2359. IEEE, 2011.

[151] Hauke Strasdat, JMM Montiel, and Andrew J Davison. Visual SLAM: Why filter? *Image and Vision Computing*, 2012.

[152] Manuel Stübler, Stephan Reuter, and Klaus Dietmayer. A continuously learning feature-based map using a Bernoulli filtering approach. In *2017 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, pages 1–6. IEEE, 2017.

[153] Sebastian Thrun, Wolfram Burgard, and Dieter Fox. *Probabilistic robotics*. MIT press, 2005.

[154] Jean-Charles Tournier and Otmar Goerlitz. Strategies to secure the IEEE 1588 protocol in digital substation automation. In *Critical Infrastructures, 2009. CRIS 2009. Fourth International Conference on*, pages 1–8. IEEE, 2009.

[155] Harry L. Van Trees. *Detection, Estimation, and Modulation Theory*. Wiley, 2001.

[156] Yanghai Tsin and Takeo Kanade. A correlation-based approach to robust point set registration. In *European conference on computer vision*, pages 558–569. Springer, 2004.

[157] u blox. u-blox announces u-blox f9 robust and versatile high precision positioning technology for industrial and automotive applications, Feb. 2018. https://www.u-blox.com/en/press-releases/u-blox-announces-u-blox-f9-robust-and-versatile-high-precision-positioning-technology.

[158] Markus Ullmann and Matthias Vögeler. Delay attacks — Implication on NTP and PTP time synchronization. In *Precision Clock Synchronization for Measurement, Control and Communication, 2009. ISPCS 2009. International Symposium on*, pages 1–6. IEEE, 2009.

[159] United States Coast Guard. GPS problem reports status. https://navcen.uscg.gov/?Do=gpsreportstatus. Accessed 2020-08-31.

[160] Roel J. P. van Bree and Christian C. J. M. Tiberius. Real-time single-frequency precise point positioning: accuracy assessment. *GPS Solutions*, 16(2):259–266, Apr 2012.

[161] Harry L Van Trees. *Detection, estimation, and modulation theory, Part I: detection, estimation, and linear modulation theory*. John Wiley & Sons, 2004.

[162] Jianguo Jack Wang, Sarath Kodagoda, and Gamini Dissanayake. Vision aided GPS/INS system for robust land vehicle navigation. In *Proceed-*

ings of the ION International Technical Meeting, pages 600–609, Savannah, GA, Sept. 2009. Institute of Navigation.

[163] Erik Ward and John Folkesson. Vehicle localization with low cost radar sensors. In *2016 IEEE Intelligent Vehicles Symposium (IV)*, pages 864–870. IEEE, 2016.

[164] K. D. Wesson, J. N. Gross, T. E. Humphreys, and B. L. Evans. GNSS signal authentication via power and distortion monitoring. *IEEE Transactions on Aerospace and Electronic Systems*, 54(2):739–754, April 2018.

[165] Kyle D. Wesson and Todd E. Humphreys. Hacking drones. *Scientific American*, 309(5):54–59, 2013.

[166] Kyle D. Wesson, Mark P. Rothlisberger, and Todd E. Humphreys. A proposed navigation message authentication implementation for civil GPS anti-spoofing. In *Proceedings of the ION GNSS Meeting*, Portland, Oregon, 2011. Institute of Navigation.

[167] Alex D. Wissner-Gross and C. E. Freer. Relativistic statistical arbitrage. *Physical Review E*, 82(5):056104–1–7, 2010.

[168] O.J. Woodman. An introduction to inertial navigation. *University of Cambridge, Computer Laboratory, Tech. Rep. UCAMCL-TR-696*, 2007.

[169] Henk Wymeersch, Gonzalo Seco-Granados, Giuseppe Destino, Davide Dardari, and Fredrik Tufvesson. 5G mmWave positioning for vehicular networks. *IEEE Wireless Communications*, 24(6):80–86, 2017.

[170] Peng Xie and Mark G Petovello. Measuring GNSS multipath distributions in urban canyon environments. *IEEE Transactions on Instrumentation and Measurement*, 64(2):366–377, 2015.

[171] Qingyu Yang, Dou An, and Wei Yu. On time desynchronization attack against IEEE 1588 protocol in power grid systems. In *Energytech, 2013 IEEE*, pages 1–5. IEEE, 2013.

[172] Haoyang Ye, Yuying Chen, and Ming Liu. Tightly coupled 3d lidar inertial odometry and mapping. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 3144–3150. IEEE, 2019.

[173] Kin S Yen, Craig Shankwitz, Bryan Newstrom, Ty A Lasky, and Bahram Ravani. Evaluation of the University of Minnesota GPS Snowplow Driver Assistance Program. Technical report, California Department of Transportation, 2015.

[174] Keisuke Yoneda, Naoya Hashimoto, Ryo Yanase, Mohammad Aldibaja, and Naoki Suganuma. Vehicle localization using 76GHz omnidirectional millimeter-wave radar for winter automated driving. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 971–977. IEEE, 2018.

[175] Thomas P. Yunck. *Global Positioning System: Theory and Applications*, volume 2, chapter 21: Orbit Determination, pages 559–592. American Institute of Aeronautics and Astronautics, Washington, D.C., 1996.

[176] Hai Tao Zhang. Performance comparison on kinematic GPS integrated with different tactical-grade IMUs. Master's thesis, The University of

Calgary, Jan. 2006.

[177] D. Zou and P. Tan. CoSLAM: Collaborative visual SLAM in dynamic environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(2):354–366, Feb 2013.